DOCUMENT RESUME

ED 048 655 EA 003 375

AUTHOR	Drewry, Galen N., Ed.
TITLE	The Instructional Process and Institutional Research. Proceedings of the Association for Institutional Research Annual Forum. (7th, Athens, Georgia, 1967.)
INSTITUTION	Association for Institutional Research.
PUB DATE	67
NOTE	171p.
AVAILABLE FROM	Dr. Patricia Wright, Office of Institutional Research, 210 T-8, University of California, Berkeley 94720 ($4.50)
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	Academic Failure, College Attendance, College Faculty, *Conference Reports, Decision Making, Effective Teaching, Grading, *Higher Education, Information Needs, *Institutional Research, Instruction, Instructional Design, *Instructional Improvement, *Instructional Programs, Teacher Attitudes, Teacher Evaluation, Year Round Schools

ABSTRACT

This conference report focuses on higher education and includes papers on the following topics: (1) the role of institutional research, (2) management information needs and decisionmaking, (3) teacher effectiveness, (4) student-faculty ratios, (5) professorial attitudes toward work environment, (6) grading and attendance in the instructional process, (7) instructional improvement, and (8) year-round education. A related document is ED 046 093. (LLR)

ED048655

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE INSTRUCTIONAL PROCESS AND INSTITUTIONAL RESEARCH

Proceedings of
Seventh Annual Forum of the
Association for Institutional Research

Edited by
Galen N. Drewry

EA 003 375

THE ASSOCIATION FOR INSTITUTIONAL RESEARCH

1967

THE ANNUAL FORUM PROCEEDINGS OF

THE ASSOCIATION FOR INSTITUTIONAL RESEARCH:

1963 -- The Role of Institutional Research in Planning[1]
1964 -- A Conceptual Framework for Institutional Research[1]
1965 -- Design and Methodology in Institutional Research[2]
1966 -- Research on Academic Input[3]
1967 -- The Instructional Process and Institutional Research[3]

1 Now out-of-print

2 Price: $2.50 in check payable to:
The Association for Institutional Research

3 Price: $4.50 in check payable to:
The Association for Institutional Research

Please request all orders from the following:

Dr. Clarence H. Bagley, Director
Office of Institutional Planning
State University of New York
College at Cortland
Cortland, New York 13045

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Association for Institutional Research TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Copyright 1967 by The Association for Institutional Research
Printed in the United States of America

TABLE OF CONTENTS

THE INSTRUCTIONAL PROCESS
AND
INSTITUTIONAL RESEARCH

PREFACE

Institutional Research -- A Future Unfolding
James R. Montgomery.............................. 1

Management Information Needs in Higher Education
Daniel D. Robinson.............................. 11

Decision Making in Higher Education
John Dale Russell.............................. 19

The Role of Institutional Research in Decision Making
Thomas R. Mason.............................. 29

Institutional Research and the Instructional Process
Stanley Ikenberry.............................. 41

FACULTY AND THE INSTRUCTIONAL PROCESS

Criteria of Effective Teaching in an Institution of Higher Education
Richard R. Perry.............................. 49

Student-Faculty Ratios: What Do They Mean?
John G. Bolin.............................. 61

Professorial Attitudes Toward Their Work Environment at the University of Michigan
Donald R. Theophilus, Jr.............................. 71

The Academic Staff Questionnaire as a Research Tool
Paul F. Mertins.................................. 79

GRADING AND ATTENDANCE IN THE INSTRUCTIONAL PROCESS

They Went to College: A Descriptive Summary of the Class of 1965
Robert J. Panos and Alexander W. Astin........... 85

Measuring Some Effects of a "Basic Ideas" Program
Joseph C. Heston................................. 95

Student Role Orientation in the Freshman Year
James V. McDowell................................ 105

The Involvement of Curricula in Student Course Changes
Bernard B. Hoffman............................... 113

Differential Failure Rates in College: Implicit Educational Concepts
Charles Bahn, Mary Cox, and Dorothy Lee.......... 117

THE INSTRUCTIONAL PROCESS

Studies of Cooperation in the Claremont Colleges: Academic Cooperation
Clifford T. Stewart.............................. 125

The Role of Institutional Research in the Improvement of Instruction
Fred Glassburner................................. 133

Research in the Instructional Process: A Model and a Proposal
Peter P. Grande.................................. 137

Institutional Objectives and Patterns of Student Change
Arthur W. Chickering.................................. 145

STUDENTS AND THE INSTRUCTIONAL PROCESS

"Year-Round Education" and the Instructional Process
Irma T. Halfter....................................... 155

Predictability of College Grades
O. W. Hascall... 157

Student Perception of Grading Practices at the University of Washington
James K. Morishima.................................... 165

An Analysis of Grading Practices
Dwain F. Peterson..................................... 169

PREFACE

This volume contains major papers read at the Seventh Annual Forum of the Association for Institutional Research. The 1967 Forum was held at the Georgia Center for Continuing Education of the University of Georgia. The Forum theme was "The Instructional Process and Institutional Research."

Space limitations made it necessary to eliminate many tables, charts, graphs and diagrams from the original papers. For readers with great interest in a particular research paper, more complete details can generally be secured directly from the authors. In order to conserve space and maintain a reasonable degree of consistency among the papers, the editor found it necessary to reduce lengthy introductions, amusing anecdotes, and personal passages. For these excisions he accepts responsibility and asks forgiveness.

The number of high quality contributions to this volume seems to be another indication of the growing maturity of institutional research. The current Proceedings reflect a greater concentration than earlier ones on the central concern of higher education -- academic programs. While finances, facilities, and enrollment projections will continue to be major fields for institutional studies, the new emphasis on instruction will provide added educational importance and balance to the domain of the institutional researcher.

Special recognition is due John Hiers, Research Assistant in the Institute of Higher Education, who ably did the preliminary editing on all manuscripts. Credit is also due typists who assisted in preparing the material: Mary Snyder, Nina Berkley, Gayle Chapman, Kathleen Coy, and Carolyn Boisky.

Galen N. Drewry
Editor

INSTITUTIONAL RESEARCH--A FUTURE UNFOLDING

James R. Montgomery
Director of Institutional Research
University of Tennessee

Man, perhaps fortunately, at this point in time cannot see into the future. But if one allowed his imagination free play, he might imagine a future campus with characteristics somewhat along the following lines.

When the incoming student arrives on campus, he reports to the administration building. As he approaches this circular structure, the doors open automatically and a pre-recorded voice asks him to enter and, if a new student, directs him to report to a designated window. There another voice asks him to place his university credit card in a slot. The card notifies a computer that this particular student has arrived; the computer in turn produces his class schedule, treasurer's statement, student permit, a prescription, and last-minute changes in registration. The prescription is for a pill to increase the individual's memory powers. The schedule is an elaborate document divided into four basic parts:

1. Traditional class sessions. The idea of face-to-face encounters between teachers and students has been hard to overcome and some regular classroom sessions still occur.

2. Learning laboratory schedule. In this schedule the student finds the times when he has access to the learning laboratory--an elaborate array of computer terminals, television and slide equipment, which is necessary for almost every course. In fact, certain courses have been replaced by the learning laboratory.

3. Motivated computerized class sessions. In these sessions a student finds each seat wired for electricity and connected with a computer. Whenever a student fails to respond or to work on assignments being presented by the computer, he receives a shock--responding keeps the current off. This type of motivation has been called negative and positive re-enforcement.

4. Library. The library remains an essential part of the learning process, although it contains an increasing amount of video tape and recorded material, as well as books. The crowds which throng it, however, make some scheduling of library time for students an absolute necessity.

While our incoming student whisks to a residence hall in a nuclear-powered campus bus, let us look briefly at other parts of the administration building. You will recall it is circular, a design necessitated by the need to place the office of institutional research (called the office of educational development) in the center, where it can be surrounded by the president, vice presidents, and related supernumeraries. Within this operation a staff of experts is busily at work. Can you guess the topics being explored?

1. That another freshman is now entering has been relayed to this office, and at the close of the day a check of the enrollment prediction will be made. This information will be entered into the SAKA (see all, know all) Model, which provides management information on housing capacity, financial needs, physical plant operations, student-bus traffic flow, learning laboratory scheduling, and related matters.

2. The staff is deeply engaged in reviewing whether this is the year that a project, delayed since 1967, should be implemented. This delayed project, as you have guessed, is on the grading practices of faculty members. The staff wonders why the same grading ratios and patterns exist as were present some twenty years previously.

There is little reason to follow this 1987 picture further. It would frighten us to find that an automatic device in the residence hall throughout the night indoctrinates, or as some say "shapes," the student in matters of school spirit, school songs, cheers, history of the institution, and subtly seeks to convince him that as an alumnus he will want to contribute financially to his alma mater. This device also instructs him on his position in the line of march in campus demonstrations, what number to call when arrested, and other such information which no self-respecting freshman should be without. Let us return, therefore, to issues which are not so far in the future and look more closely at them.

Institutional research, whether by that name or another, appears to be a part of the campus scene for the foreseeable future. With approximately twenty new colleges opening each year and with enrollments increasing in most of the older ones,[1] expansion of higher education in general and institutional research in particular appears assured. The variety of students with the accompanying accumulation of experiences, knowledge, and interests will continue to bring problems. More federal money for higher education will cause so-called private as well as public institutions to develop more extensive record-keeping systems, while the need to develop more evaluative programs and procedures will be obvious.

Two groups increasingly will want information. The administration, faced with operating within a bureaucracy, looks across an institution where, with rare exception, each department takes an active interest only in its own affairs. Faculties and trustees emotionally want strong leadership,[2] although they are somewhat less than enthusiastic about the person who demonstrates too much of it. The point is that an administration, if it is to move forward, needs help, and to the extent it can be persuaded to use institutional research in making studies, the task will be less arduous. The second group, which in the future will want more information, is the faculty. In complex institutions the person who has the figures and knows what they mean is a formidable opponent for a less well-informed individual or group to debate. With faculties already showing a tendency to move toward collective bargaining and stronger senate activities, and showing little inclination to drop traditional suspicions of administrative behavior, the faculty will want more and more information. If institutional research fails to supply it, other channels of communications, perhaps less appropriate, will be developed.

In order to meet the needs of both the administration and the faculty, therefore, institutional research must undertake assignments in two vital areas: operations research and academic surveys. To the extent that we would like our efforts in these areas to have an impact on the institutions we serve we should not forget that for the past two years our keynote speakers (Allan Cartter and Lewis Mayhew[3]) have harangued us over the lack of impact of institutional research. We must not assume that our presence on campus and the weight of our reports automatically insure that our work is "functional"--will influence in any substantive way the decision making process on our respective campuses.

Two Basic Types of Studies for the Future

With rare exception every college and university, from the smallest to the largest, needs two basic types of studies: operations-management research and academic development research.[4]

1. Operations-management research might include such factors as the following:
    a. Long-range and medium-range planning for campus physical development.
    b. Realistic cost studies, space studies, and similar reviews which might help to stretch dollars, especially in light of rising costs.
    c. Simulation studies and models for management operations.

2 Academic development research might encompass such studies as the following:

a. Long-range and medium-range planning for academic development.

b. Learning studies.

c. Value studies.

Studies in Operations Research

The need for long-range planning is obvious to one in institutional research. Certainly questions concerning operations research should be included in it. For example, to what extent is planning, with or without research, undertaken in developing better communication and communication skills on college campuses among the various groups and individuals?

The Rourke and Brooks report notes the failure of college administrators to introduce major breakthroughs in what might be called scientific management. Neither institutional research officers nor computer managers have thus far brought extensive change.[5] Yet we know that a financial study, when conducted by a man such as John Dale Russell, may cut to the center of a university, clearly showing defects and remedies.

Many cost studies serve more for eye appeal and fad appeal than for practical management purposes. There are real exceptions to this generalization. Such studies, regardless of their external purposes, would seem as a minimum to allow decisions to be made, with some awareness of the outcome, on how much could be allocated for raises, for expansion of faculty, for new operations, or for additions to departments which have the greatest need for personnel.

Prediction and simulation research frequently lead to areas which contain short-term information of great value. At The University of Tennessee this year an effort to predict enrollment by courses had several immediate and conflicting reactions on our campus: one department head and one dean thought we were too low, one vice president thought the material helpful, another disagreed with our beginning assumptions (with partial correctness), one department head thought us too high but nevertheless adjusted his figures accordingly, and another believed us correct. Hence, this report for The University of Tennessee, when further refined, may prove a helpful management tool. Our institutional research people borrowed the idea from another institutional research office[6] and introduced it--top management did not request it.

Perhaps the most involved model of a university operation is one under study by H. E. Koenig and associates at Michigan State University, developed with the assistance of a grant from the National Science Foundation.[7] The model builders want to observe the relation between flow of students and resources. The model seeks to identify the demands of the students on the institution and the resources needed to meet them. The model, admittedly, is far from perfect, and its complexities have not been solved completely. But it represents a major effort to enter into model building which might aid administrators.[8]

In operations research it too frequently happens that one becomes convinced that a computer on campus is necessary before undertaking research. Nothing could be further from the truth. The third generation computers with time-sharing apparatus should enable any college with an enrollment of five hundred or more to obtain assistance from a nearby installation. Furthermore, it is possible to spot an enrollment trend, the need for a residence hall, or the need for changing a master scheduling plan with a calculator or simply a pencil.

It is not necessary to wait for a grand breakthrough. For example, Koenig's model may work to develop smaller, less complex models or other types of predictive information. L. E. Hull and associates at Indiana University are well launched in combining less complex simulation studies which work into helpful management information materials. Although a campus lacks a total information system, there is nothing to prevent what information is available, whether computerized or not, from being utilized effectively. One should be able to predict the effect of converting a classroom to office space, to review a telephone system, to ascertain the desirability of a larger or smaller summer school session, or to know whether added faculty members go to departments with the most need (at least from a student-served basis), regardless of whether a total information system exists or an elaborate model is available.

Frankly, top level administrative officials frequently cannot define the precise information they need. At times institutional researchers are not aware of basic problems but instead receive requests which someone hopes will produce a helpful report. We have undertaken in our office too many descriptive studies which had length but no impact. Sometimes it is the fault of the design; sometimes of the initial question. Two involved space utilization studies grace our shelves. These show in great detail the percentage of rooms utilized on a forty-four hour week basis, the percentage of student stations utilized, and the percentage of use in comparison with other institutions. How many more students could be accommodated, how many sections could be added before reaching capacity, how long before a twelve or fifteen period class day appeared necessary, what was the relative need between classroom and office space, and a host of other unidentified items were perhaps needed more.

Operations research for institutional researchers means supplying the right data, in the correct amount, at the appropriate time to aid in policy making or development procedures. Regardless of institutional size or data processing equipment available, the objective is first to identify the basic problem and then to find the applicable answers related to it. Such research should be backed with theory and related whenever possible to academic surveys (or development).[9]

While operations research presents one aspect of the task in the foreseeable future for institutional researchers, academic surveys represent another area which only now is beginning to emerge and attract the attention of research offices. This field of research holds as much, if not more, promise of producing an impact than operations research.

Studies in Academic Development Research

As a starting point for obtaining better information on faculty interests, a survey might be in order to find problems with which faculty members want assistance. Perhaps to identify, for example, an interest in receiving help with test construction or computer assisted instruction procedures would reveal potential areas for assistance and for research.

Necessarily, studies of grading procedures would fall within this focus. Presently there is a need for an extensive monograph on grading practices in colleges and universities. Such a report might collect studies, such as those of John Hills[10], review findings, and begin to ascertain the theory which pertains to this relic of higher education. Grading practices may prove to be the most important variable in studying learning and retention.

Most institutional researchers have neglected meaningful studies on curricula and courses. Dressel's frontal attacks against proliferation of courses and small course enrollment which serve to weaken the total educational structure represent one effort.[11] More study of a similar nature is needed, as well as of other potential areas, such as the short-term and long-term relevance of taking five or six courses per term rather than three.

Another area needing attention might be called the learning environment. How one learns, from the general to the specific or specific to general, and the factors involved which cause a person to retain and transfer or apply material learned are educational unknowns which lie within our research grasp. This is not to suggest more studies on "how to teach effectively" by lecture, television, discussion, or similar techniques. Rather, it is time for teachers to know the goals toward which they are striving and the persons whom they are teaching. Assuming subject-matter competence, one might explore with teachers:

1. Ways to organize course content to encourage learning regardless of whether they disrupted the traditional classroom routine.

2. Ways to teach a class so that each of the diverse elements within it may profit from the material, or stated simply: to challenge the assumption that the entire class must be taught in the same manner.[12]

What evidence really exists to support contentions that students learn to think and respect honest difference of opinions, both highly proclaimed end products of higher education? What differences do teaching arrangements make (for example, what is the evidence to demonstrate that the same amount of class time is needed for English as for history, or mathematics as for zoology)?[13] What evidence do we find that the college or university has had an impact on the student, and if so, did this impact have anything to do with the classroom presentations? Can students with similar learning characteristics be identified and placed in similar learning environments?[14] What is the evidence to support the desirability of developing cluster colleges? Finally, what might happen if we asked the students how they wanted to learn and tried teaching accordingly?

Already computer assisted instruction is a reality at a few universities, and within the next few years development of this learning procedure will accelerate. Institutional research needs to be a part of this development and ready to look at such factors as effectiveness and costs. In addition, it might be appropriate to begin to explore the "proper" division between computer-aided and teacher-aided instruction. Only a step away is "package curricula" from corporations which have combined into a packet materials such as printed matter, audio-visual tapes, or items to observe and smell.[15] While the market for such material may not initially be found in higher education, evaluation of such products should be in order.

One final illustrative area for academic research might be value development. While some of our institutional research members have worked within this field, we all need to be familiar with the literature and research implications of it. Junius A. Davis', "What College Teachers Value in Students," should cause us continuing worry. Davis asked entering students at a high-prestige engineering college to describe themselves on such factors as interest in ideas or devotion to scholarship. Three years later faculty members were asked to select the students who exemplified the highest goals of the institution and who appeared likely to excel as engineers. As you have already guessed the descriptions from the students and the perceptions of the faculty lead one to wonder about the values of both groups. Those students describing themselves as talkative and socially confident had in fact emerged as "well known" to the faculty. Those students thought to be heading for successful engineering careers had pictured themselves as preferring newspapers to books, as being "calm and deliberate" rather than "venturesome," and as preferring to use a computer rather than to design one.[16] The implication for institutional researchers is not necessarily whether these values are bad or good, but whether our offices have the faintest perception of student and faculty values.

Studies concerning learning are difficult to undertake for a variety of reasons (including faculty indifference), but this difficulty does not negate their importance. The field awaits and needs institutional research personnel either to undertake or assist others in studying this emerging area of concern.

Enhancing the Likelihood of Impact

There are several ways in which we might increase the influence of our efforts in coping with formidable research requirements and have some impact on the policies and programs of our institutions. There are five essential elements which should become the *sine qua non* of institutional researchers if our image is to become what we desire: we need to write more, and more lucidly; we need to develop campus-wide stature for honesty in reporting findings; we need to become more concerned with theory and with experimental design; we need to evaluate our effort; and we must say what we believe the implications of our research to be for policy formulation.

Let us look first at the topic of writing, especially for readers outside our own institution. The most recent book published on institutional research and operations research is the *Managerial Revolution in Higher Education* by Rourke and Brooks. Neither of these men is an institutional researcher, at least in title, but this book makes a helpful contribution to this field. The *Annotated Bibliography* of the Association, first published this year, is a step in the direction of letting others know what is being written; moreover, it will allow some idea to be formed of work completed or underway. When one serves in an institutional research capacity, it becomes obvious why writing is difficult: the press of assignments and the unspectacular nature of most resulting reports make writing, for other than limited internal use, difficult. But the problem may be more basic: with occasional exceptions most people need a strong incentive to write, such as the desire for a promotion, a raise in pay, or the applause of other members of an academic discipline. These incentives, such as academic rank, are frequently less strong or more removed for us.

A second tack is to develop a campus-wide stature for honesty in reporting findings. To fudge data slightly, to make a one-tailed statistical test when a two-tailed one is in order, or to take any short cuts in seeking the correct answer must be abhorred. To be honest and to give all the data may be extremely difficult--especially when the president or some other individual has expressed his idea of the direction the study should go, or at least what he expects in the recommendations. Paul Dressel summed up the matter extremely well in a recent speech:

> I think an institutional research director should be too honest and forthright to be a college president, and he should be too much interested in the improvement of higher education to waste time being a dean. He should be more interested in focusing attention on problems and getting people to come to grips with them than in being loved. He should see himself more as a faculty member than as administrator, but be willing to recognize that faculty members will regard him as an administrator and administrators will see him as just a little bit too oriented to the faculty point of view to be completely trusted.[17]

Until one has turned in a study only to find that a decision has been made in a direction other than that indicated in the study, one has not been initiated into the interfraternity of institutional researchers. If recommendations and findings from institutional research officers are always in the direction the administration or faculty wants, then we are failing our responsibility.

A third area of continued emphasis should be on research methodology. The 1965 Forum focused on the theme "Design and Methodology in Institutional Research".

Although we recognize the necessity of using theory or of developing theory for our studies, frequently it is either de-emphasized or disregarded. Kenneth M. Wilson's reports from the College Research Center in Poughkeepsie represent a notable exception,[18] and there are other researchers who also make this effort.

In addition to tying theory and related research more closely into surveys, there remains much to be undertaken with experimental research designs. The Campbell and Stanley chapter in Handbook of Research on Teaching[19] should be familiar to each of us. Yet, the material emanating from institutional research offices is almost completely devoid of investigations that have been based upon experimental designs. Unfortunately, some descriptive studies, which are so laboriously prepared, contribute no more than a news release--and sometimes even the press-release personnel cannot find any item of value to report. The "Fact Book" of The University of Tennessee, which has required some six months to complete, is an excellent case in point, for after it is distributed this month it will be carefully filed and forgotten. The same amount of labor and investment in an experimental study might have been spent more profitably.

On the other hand, experimental studies are neither easy to start nor to finish. Personnel from our Office have sought permission from appropriate administrative officials to undertake six experimental studies in the past two years. Five of the six requests brought quick refusals, but one on a topic of learning improvement holds great promise. Even with administrative backing, however, such studies on occasion come to grief. At the Forum last year, James Morishima reported on an experimental research study which grouped certain academic majors within residence halls. This study became the object of a student power struggle which produced resistance to participation in an activity in which purposes and procedures were not clearly known by the students[20]--a contention with merit but one which would certainly contaminate the control and experimental groups.

It is also past time to evaluate our efforts as institutional research officers. To assess where we are and possible directions for movement, it appears to be time for surveys of faculty attitudes toward institutional research endeavors, and some evaluation of the extent to which our findings are getting across to faculty and administration. The attitudes which people hold toward institutional research should be of interest, but even more important their ideas on studies and approaches should be revealing. This year we tried a small experiment along these lines with rather disastrous results: instead of placing an evaluation form at the front of a report of one of the studies completed by the office, that we might collect a reaction to the study, zip code information, and suggestions for further studies, we placed it as the last page. Out of a mailing to sixty on-campus and off-campus individuals, ten of the evaluation sheets came back to us. Hence, our reports do not command the reading appeal of a spy thriller. We plan to change our format for reports--whether people will read them remains to be seen.

The fifth element concerns the need to press for policy decisions based on survey findings. A recent survey of institutional research officers indicates that less than half of them submitted recommendations with their studies,[21] while identifying directions pointed out in a survey and stating recommendations appear essential elements for a good report. The information which appears so clear to us may not have a similar meaning to the layman. When a report is submitted and one self-righteously steps aside to let another find the meaning, it should not be surprising if nothing is undertaken or the reverse of our intention occurs--and we have no one to blame but ourselves. Perhaps this is one reason why Rourke and Brooks report that college presidents placed a lower estimate on the influence of institutional research in policy making than institutional research directors believed they had.[22]

The future for institutional research appears bright. But before we reach the center of administration building or even remain on a college campus, we will need to produce better and more extensive writing, to be honest always with our surveys, to develop better methodology, to evaluate our efforts, and to take stands which may influence policy making. While standing for these essential elements for institutional research, our studies will need to be concentrated in the area of operations research and academic development.

FOOTNOTES

1. Frank Bowles, Wall Street Journal, February 13, 1967.

2. Nevitt Sanford, "Leadership for Improved Conditions for Learning and Research," (mimeographed, Association for Higher Education, 1967), pp. 3-5.

3. For example see Lewis B. Mayhew, "Imperatives for Institutional Research," Clarence H. Bagley, ed., Research on Academic Input (The Association for Institutional Research, 1966), pp. 1-7.

4. These categories are discussed in some detail in Francis E. Rourke and Glenn E. Brooks, The Managerial Revolution in Higher Education (Baltimore: Johns Hopkins Press, 1966).

5. Ibid., 104.

6. Indiana University.

7. H. E. Koenig, et al., "A Systems Approach to Higher Education," Interim Report No. 3, Project C-396, National Science Foundation (Washington, D. C. and Michigan State University, 1966).

8. Other efforts at model building are being tried at such institutions as the University of California, Maryland, and Purdue. Rourke and Brooks, Managerial Revolution, p. 32.

9. This is discussed more fully in Henry S. Dyer, "Can Institutional Research Lead to a Science of Institutions?" The Educational Record, XLVII, No. 4 (Fall, 1966), pp. 452-466.

10. For example see John Hills, "The Study of Factors Influencing College Grading Standards," (Research Bulletin 2-66, Office of Testing and Guidance, Board of Regents, University System of Georgia, 1966).

11. For a recent article on this topic see Paul Dressel, "Specific Points of Attack in Curriculum and Course Revision," Journal of Educational Research, LIX, No. 7 (March, 1966), p. 314.

12. Laurence Siegel, "The Contributions and Implications of Recent Research Related to Improving Teaching and Learning," in Ohmer Milton and Edward Shoben, Jr., eds., Learning and the Professor (University of Tennessee, Knoxville; Learning Resources Center, 1966), p. 132.

13. Ohmer Milton, "The State of the Establishment," Learning and the Professor (University of Tennessee, Knoxville; Learning Resources Center, 1966), p. 7.

14. Siegel, "Contribution," p. 143.

15. Report from Education Summary, January 1, 1967, cited in Education Recaps, VI, No. 5 (February, 1967), p. 11.

16. Junius A. Davis, "What College Teachers Value in Students," College Board Review, LVI (Spring, 1965), pp. 15-18.

17. Paul Dressel, "The Nature of Institutional Research," (1966), pp. 7-8.

18. For example see Kenneth M. Wilson, "Assessment of the Graduate Study and Career Plans of Seniors at Three Liberal Arts Colleges for Women: A Pilot Project," (Poughkeepsie, N. Y.: College Research Center, 1967).

19. Donald T. Campbell and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," in N. L. Gage, ed., Handbook of Research on Teaching (Chicago, 1963), pp. 171-246.

20. James K. Morishima, "Effects on Student Achievement of Residence Hall Groupings Based on Academic Majors," Bagley, ed., Research on Academic Input, pp. 163-170.

21. Rourke and Brooks, Managerial Revolution, p. 63.

22. Ibid., p. 60.

10/11

MANAGEMENT INFORMATION NEEDS IN HIGHER EDUCATION

Daniel D. Robinson
Partner
Peat, Marwick, Mitchell and Company

The U. S. Office of Education recently published a booklet entitled "Projections of Educational Statistics to 1975-76" which provides a succinct picture of the growth of higher education in the United States. The report covers the actual statistics for the ten year period 1955-56 through 1965-66 and estimates for the next ten years ending 1975-76. A glance at some of the figures for institutions of higher education yields a perspective for this period of growth.

	Actual 1955-56	Actual 1965-66	Projected 1975-76
Fall degree-credit enrollment	2.7 mil.	5.5 mil.	9.0 mil.
Earned degrees - bachelor's through doctor's	378,000	680,000	1,202,000
Full time equivalent instructional staff for resident degree-credit courses	161,000	306,000	454,000
Total expenditures	$5.0 bil.	$15.2 bil.	$25.3 bil. (1965-66 dollars)

These projections involve a number of assumptions as to population growth, continuance of trend lines and so on. Most projections made over the last decade have proved to be conservative. Hence, it is safe to say that higher education in this country is reaching immense proportions. And as a result, the entire character of higher education has been changing and, in all probability, will change more rapidly and drastically in the ensuing decade.

The size and growth of our system of higher education has placed an ever increasing burden upon those who are concerned with its governance. In earlier years, particularly those preceding World War II, the typical college or university was characterized as having a "relaxed academic atmosphere." While there certainly were problems to be faced and overcome, the pace of institutional life was slower, less complicated, and less demanding from an administrative point of view. The institutions were smaller, their programs less complex, their growth less dramatic.

In the decade 1930 to 1940, resident degree-credit enrollment increased by a total of 35 per cent. The average enrollment of all institutions of higher education rose from 780 to 875 students in the same period. By 1964, the average enrollment was more than 2,300 students. Even more illuminating is that by 1964 almost 90 per cent of all degree-credit students were enrolled in less than half of the number of institutions having 1,000 or more students. Furthermore, over 40 per cent of the students attended institutions having enrollments of 10,000 or more, which were less than 5 per cent of the number of institutions. Thus, the emergence of the large university is part of the pattern of change. So, too, is the shift in proportional enrollments from private to publicly supported institutions. The growth in research and development expenditures, the increasingly bewildering array of federally sponsored programs, the importation of foreign students and the exportation of faculty and technical staff to foreign lands add to the urgent demand for better, more effective administration. No longer may college and university presidents make horseback guesses, operating by the seat of their pants, in the hope that everything will come out all right in the end. The war on ignorance and human obsolescence in an increasingly automated, technological society is a deadly serious endeavor and will bring with it institutional casualties whenever management fails in its mission.

If educational administrators are to meet this challenge and act intelligently and purposefully, then they need information--which is the main topic under review here. But before an attempt is made to identify the information needs of administrators, we must first decide who they are. Recently the American Association of University Professors, the American Council on Education, and the Association of Governing Boards of Universities and Colleges formulated a pronouncement on this subject entitled, "Statement on Government of Colleges and Universities." This document is concerned with the divergent and interdependent roles of the governing board, the administrators, the faculty and the students. While some will argue with its conclusions, and others will be frustrated by its omissions, one fact is perfectly clear. The way an educational institution is administered is far different from the mode of the relatively monolithic, pyramidal structure of private industry which educational administrators so often attempt to emulate. The lines of authority are not clear-cut, the willingness of individuals to assume responsibility is not uniform, and the interests of each group are inextricably woven together.

Even in the area of what might be called official administration, there are unique problems. Most students of college and university management have increased the emphasis on the critical role of the academic department chairman as the first line of management. Yet, how does the training, the experience, or the personal inclination of the man equip him for his managerial post? Many department chairmen are selected on the basis of their scholarly attainment, public recognition, teaching ability, and so on. Not a few accept the chairmanship reluctantly, looking forward to the time when they may drop their administrative duties and resume full-time teaching and research activities. While all of this is understandable, it can hardly be accepted as proper preparation for the demands of administrative functions. One can move on up the ladder of the academic hierarchy and find many of the same characteristics.

The typical governing board has at least some members who have been trained in, and have successfully executed, management responsibilities; but most have gained their experience in industry or government. The uniqueness of higher education makes them feel uncomfortable, and they tend to focus on familiarities like financial budgets and investment portfolios. Perhaps this explains the rather limited role which the previously cited policy statement spells out for the governing board. The one fallacy in this defined role is that an uninformed governing board is seldom an effective mechanism for raising funds from either private or public sources. If the role of the governing board in securing the ever increasing financial resources is to be sustained, some change must take place in its involvement in institutional affairs.

The faculty in most institutions exercises control over many critical facets of institutional development and operation. Curriculum development, degree requirements, admissions standards, and research policies are all areas of faculty involvement, if not outright control. When the implications of decisions in these areas are understood, the vital effect they have on the use of resources can be seen. And, while the faculty has important authoritative prerogatives, it seldom has the responsibility of finding the required resources.

Presently, this effort to define who management is has not been very illuminating. Perhaps there never will be a simple, straightforward answer. Perhaps there never should be one. As long as this confusion exists, however, the prime requisite of an effective information system will be missing, and that is to define who will receive the information in the first place.

Let us assume now that "the administration" has been defined in our college or university in some finite way. There still remains another hurdle before we can talk about information: who makes what decisions? A basic postulate of the function of information is that it is of assistance to the decision making process of management. It must be possible, therefore, to identify not only who must make administrative decisions, but also what kinds of decisions a given administrator must make and what kinds of activities they affect.

Once again, the problems inherent in the nature of higher education become apparent. To the extent that authority and responsibility do not go together or that different groups or individuals have overlapping concerns, it becomes increasingly difficult to find the point in an academic institution at which a final decision is made. The extensive use of committees to cope with a variety of problems and policies is one of the hallmarks of the academic community. While this practice has a long history and is widely used, it can pose problems in fixing final responsibility for decision making. Some have said that the very act of decision making is anathema to the academic mind which is nurtured in the spirit of inquiry, contemplation, and discussion. These traits are fundamental to good teaching and imaginative research, but they lack the element of decisiveness which ultimately must characterize managerial action. When these obstacles to decision making are recognized, the often heard phrase "university policy" is understood more clearly as the mask which covers the unknown decision maker.

As our administration has been defined by assumption earlier, let us also assume that responsibility for decision making has been established as well. Now we come to the heart of the matter--management information for decision making. But remember that the heart cannot function without the head.

If any area of institutional activity has come into its own, it is that of the collection of data. The chief culprit in this mushrooming exercise is the computer. As each new generation of computers becomes operational, the ultimate goal seems to be more and more possible; namely, to accumulate all information about everything. A great deal of fun can be poked at computers and those who live with them and by them, yet few can visualize how the myriad transactions and activities of our numerous institutions could be coped with in their absence. The emphasis in computer applications in the past has been on what some refer to as "housekeeping" activities. The accumulation and manipulation of data for routine clerical activities which has been the mainstay of registrars' offices, accounting departments and the like, have been automated to a high degree in many large institutions. The production of informational reports as a part of this routine processing has been considered a by-product. It is the area of management information which currently is receiving widespread attention and ultimately will prove to be the most important contribution of computer technology.

However, the trouble with computers is that they will only do what they are programmed to do by people, and too often the people do not make very intelligent decisions. The easiest way out is to produce reams of reports, at a thousand lines a minute, which literally overwhelm the people who receive them with a mass of undigested, irrelevant data. The capturing of raw data and the translation of it into machine-readable form is an art which is progressing faster than the ability to produce the proper kinds and amounts of management information. While it would be a waste of time to define all the different kinds of specific information which administrators might use to good advantage, it is essential that this type of information be characterized generally. Criteria for the evaluation of management information must be established first. The specific content will flow naturally from this definition.

To gain perspective on the problem of criteria, it may be helpful to consider some concepts developed by a committee of the American Accounting Association, as set forth in its Statement of Basic Accounting Theory, published in July, 1966. This rather forward-looking statement suggests four basic standards to be used in evaluating accounting data: relevance, verifiability, freedom from bias, and quantifiability. While adherence to all four standards is not always equally possible, each should be weighed and considered. Fortunately this statement exhibits a concern for statistical as well as financial data; hence, the applicability of its conclusions to the affairs of higher education is relatively easy to imagine.

As previously stated, much more information about many aspects of institutional activity is available now than ever before. The first problem in adapting the data to management needs is to discover how it relates to both the area of activity being administered and the type of decision being made. Most activities in a college or university generate numerous individual transactions, yet few people need, or could absorb, each transaction separately. Summarization of individual transaction or activity data must be employed so that the user of the information can see patterns and structures emerging from the mass of detail. If the data becomes too highly summarized or generalized, it loses its significance again. What must be found is that level between absolute detail and absolute summarization which has the highest degree of correlation to the problem at hand.

When we speak of moving from minute lists of specific information to the generalized summary, we are describing an ordered leveling of informational groupings. This hierarchical structuring of information is the most critical element in classifying data in an information system. Those who have gone through the pangs of giving birth to a computerized total information system, or even to the development of a portion of it (such as a chart of accounts) have experienced the agony of laying out an orderly classification system which will provide a logical structuring of the detailed information to be processed. In effect, the administrator who uses information for decision making wants to screen that which is irrelevant and consider that which is relevant. What he finally considers must then be assembled in digestible form without losing its integrity.

Superficially, it seems simple to say that you accept the relevant and eliminate the irrelevant. Yet problems arise not only in the choice between detail and summary information in the same area of activity, but also in the interrelationship between many different types of activities. Whenever a decision of any magnitude is made with respect to one aspect of a college's operations, it invariably has a direct effect on most other activities. For example, an admissions policy may be developed after searching inquiry respecting test scores of incoming freshmen, class standing in high school, rate of attrition in college, number of freshmen on probation, and other similar information for a given period of time. Once an admissions policy is established and instituted, however, it has far reaching effects on such matters as faculty, space utilization, costs of instruction, and residence hall demands. Thus, we see that while a decision can be made in light of knowledge of the circumstances surrounding the subject matter of the activity itself, the effect of that decision on other activities must be considered as well. The reactions which take place in these interrelated activities, if known in advance, could lead to a totally different decision. The universe of information which is relevant to a given decision tends to expand as the results of the implementation of that decision are traced through all the affected parts of the institution.

Once data is assembled and is shown to be relevant to the subject at hand, the administrator may be troubled by the second and third of the four basic standards mentioned above: verifiability and freedom from bias. As educational institutions grow larger and more complex, the president and his chief associates get further away from the grass roots of daily activity. More reliance must be placed by top administrators on what they are told by subordinates. The computer once again enters the picture as a symbol of depersonalization, remoteness, and excessive size. But the computer is the effect in this case and not the cause; it stands between the originators of detailed data and the user of its end product reports. The user-administrator can easily become suspicious of the computer, its operators, and his subordinates, and he will seek means of ascertaining the reasonableness of the data supplied to him. Most experienced administrators intuitively seek for another way at arriving at facts and figures which will tend to substantiate the reliability of the information first supplied.

Bias in information is very subtle and sometimes disastrous. Careful choice of measurements or indices used to express quantified data is always a problem. The object, of course, is to eliminate a preconceived objective in the manner in which the subject matter is being measured. Take, for example, the glass containing a certain quantity of water. One might say that the glass is half full, another could state that the glass is half empty. Both would be correct, yet a more unbiased statement would be that the glass held fifty per cent of its capacity. It is to be expected that in any large organization where a number of men participate at high levels of responsibility, albeit in different areas of endeavor, each will develop strong feelings about the institution as a whole. Through this sense of zeal and well motivated enthusiasm, reports of seemingly realistic data can be presented in such a way as to support an inescapable decision which aims the institution in a direction consonant with the predilections of the reporter.

Another type of bias problem occurs which is equally dangerous and which is an effect of the computer. Prior to centralized data processing, certain types of data were the domain of separate administrators. Student records were maintained manually by the registrar and his staff. If anyone wanted to know anything about the academic record of these students, he had to get the information from the registrar. Similarly, the controller and his staff kept books by hand or by means of locally controlled equipment. All financial reports had to be obtained through the controller or with his permission. Now, all of this information and much more can be maintained in one place-- in the highly impersonal memory of a computer. Now, if he wishes, the president can get any and all of this data from the computer without referring to the registrar, the controller, or anyone else except the data processing director. While this short-cut to information is enticing and may even be freer from one type of bias at times, it leaves out the very important element of competence in the area of the information itself. The function of the registrar, the controller, the director of admissions, et al, is not simply to collect data, but also to interpret it in many respects. When this vital human element is eliminated, the information can become next to useless, if not misleading.

The problem of the fourth basic standard--quantifiability--reaches its zenith in institutions of higher education. The very terms that are closest to the heart of academe so far have eluded quantification. Academic excellence, quality of instruction, liberal education, adequacy of research necessary to instruction, and the extent of participation in the affairs and problems of the community are all basic to the goals and aspirations of the vast majority of colleges and universities. No one has the temerity to say that he has the formula for converting these characteristics into precise numerical values. On the other hand, every educational administrator readily makes decisions regarding policies which he feels are expressions of those qualities. The establishment of a particular class size in a given course of study in a particular discipline, the prerequisite of advanced degrees for promotion to professional ranks, the selection of students on the basis of test scores and class standing, and a host of other policy decisions which are made every day are really quantifications of a predetermined sense of quality or mission. The problem, however, is that in too many cases, the quantification in the form of adopted policy is not clearly related to the particular quality which is sought. How often have admissions standards been raised by institutions, only to discover later that the rate of attrition has not dropped appreciably? How many successful governing board members and alumni have complained that they could not have attended their alma mater in the face of today's entrance requirements? How many times have outstanding men in industry, government, the arts, and teaching been cited as only mediocre students? Despite these contradictions and in the face of the seemingly hopeless struggle to measure such abstractions as quality, efforts must be made to improve these techniques or else the whole mission of providing information for decision making fails.

So far, we have dealt with the problems inherent in providing information to academic administrators. Perhaps we should turn to the brighter side for a moment. The most significant improvement in academic administration in recent years is the growing awareness of the need of more information about the institution to aid the administration. This awareness has led to an impatience to begin the business of establishing a broad base of data and the means of manipulating it for useful purposes. In turn, more resources have been allocated to the job of developing this capability in a growing number of colleges and universities. Most significantly, the field of institutional research has been given support and encouragement to play an even more significant role in support of academic administration. The sense of inadequacy that grips the administrator of today urges him to look for better qualified staff and more sophisticated managerial tools to assist the administrator in his role of directing and controlling the institution's human and material resources.

One of the more advanced managerial techniques which has intrigued institutional researchers and others in higher education is the computer model. The purpose of the model is to express mathematically the myriad interrelationships that exist among the elements that constitute a part or all of the structure of a college or university. Once properly constructed, the model can then simulate the manner in which the institution utilizes its resources under various assumed instructional, research and public service demands. Such a model would enable the administrator to anticipate and understand the effect on all areas of activity of a policy decision in one activity or function. The model could be used to simulate a wide range of alternative decisions, providing a choice which would result in maximum overall benefit. The same model could also provide the long-range effect of today's decisions. As conditions change and, with them, the interrelationships of which the model is formed, the model itself could be updated and made to conform to reality on a current basis. Visualize a loop which starts with the model constructed on the basis of immediate past performance, then used for projection, then updated by new information taken from ongoing actual experience, then used in its updated form for new projections.

The computer model could bring another popular technique within reach of the academic administrator. This technique, popularly called the planning-programming-budgeting-system or PPBS, has enjoyed immense popularity in defense industry and is being applied in various ways to many other types of industrial and governmental activity. The technique involves five major phases: (1) identifying goals, (2) establishing order of priority, (3) determining and selecting from alternative courses of action, (4) projecting the impact of current decisions on future years' operations, and (5) measuring performance against the plan. While several of these steps do not necessarily require a computer model, those involving selection of alternatives and projection over time would not be effective or realistically feasible without the computer model or some adaptation of it.

Let us finally turn our attention to what must be done in the future if the needs of administration for useful information for decision making are to be met. A review of the problems mentioned earlier suggests the answers. Administrative responsibility and authority in academic institutions must be studied more carefully and clarified, without being made so rigid as to be destructive of the essence of the institution. Information must not only be gathered, but also organized and structured so that it is relevant to the decision to which it is directed. New relationships between cause and effect of alternative policy decisions must be found which will give clearer insight into the reasons why things happen in these institutions. Better measurements based on scientifically proven studies must be developed which will enable a more reliable quantification of quality factors. More realistic alternatives for program execution must be developed through analysis of significant samples from a wide variety of practices.

The role of institutional research in the next decade can be the deciding factor in the race to meet the demands of our society for more education for more people than any other country in the history of the world has ever attempted to provide. Whatever is accomplished must be done within the confines of limited resources, and these constraints will become more noticeable over the years ahead. The competing demands of our society for all manner of social and physical improvement will force a limitation on the money and staff which will be available to higher education. Research must find the answers to the many problems which will otherwise stand between effective, relevant education and the citizenry it is intended to serve.

18/19

DECISION MAKING IN HIGHER EDUCATION

John Dale Russell
Consultant, Office of Institutional Research
Indiana University

Perhaps the term "decision making" has an important advantage over the older terms, in that it tends to focus attention on the process of administration, rather than on its structure. I shall interpret the term rather literally and shall try to limit discussion to those aspects of administration or control that actually do involve the making of decisions.

What is a decision? The most suitable definition given by Webster is: "Act of determining in one's own mind upon an opinion or a course of action." Note that a decision is a mental act, a rational act. Note also that the primary unit in a decision is an individual person. This is an extremely important part of the concept implied in the term "decision making." A decision is sometimes reported by a group of persons, such as a committee, but such a decision is always the product of the minds of the individual persons in the group. The decision may involve only an estimate of the balance among individual, personal decisions; for example, a new rule may be promulgated because some official estimates that "this is what the students want." Or, the decision of the group may be just a consensus, meaning that no one is sufficiently concerned to object to what the leader or leaders of the group have decided. Often the decision of a group involves a count to determine what the majority of the members have decided. But in every case such decisions are the product of individual decisions by the members of the group.

That a decision is a product of the human mind, and is therefore made by some individual person, leads to two important corollaries. It is in most cases important to know who made the decision, and that the decision-maker had authority to make the decision. The precise location of the source of a decision is essential in good administration. If the decision is by a group of persons, such as a committee, it is sometimes convenient for some of its members to hide behind the anonymity of the group. But unless one is on record with an objection or a contrary opinion or recommendation, it may always be assumed that he shares full responsibility.

It may be useful to classify decisions for further analysis of the decision-making process. There are at least four different systems of classifying decisions that have to be made in higher education. Perhaps the most customary method places them in two groups: those that are large or important, and those that are small and relatively unimportant.

In a well administered institution one would expect to find that large or important decisions are made finally at a high level in the organizational structure, and the small decisions are made at the lower echelons of executive responsibility. For example, the decision to add a new school or college to the university organization would typically be made at the level of the president, board of trustees, and sometimes even the state legislature. But action on the request of a janitor to take a day off from work to attend the funeral of his wife's third cousin would ordinarily be taken by a foreman or supervisor several echelons below the president's office. The rule on the assignment of important and unimportant decision-making responsibilities is, however, often disregarded.

But, one difficulty in classifying decisions into small and large arises from inability to foresee their long-term effects. What may seem a particularly difficult and large decision at the moment may, in the long run, have no important consequences for the institution. By contrast, a decision that for the moment may seem inconsequential may ultimately prove to have important effects on the whole future of the institution's program. Two of the most important kinds of decisions ever made in a college or university often fail to get the high level of attention they deserve: the appointment of any new member to the faculty, and the granting of tenure to a faculty member. The sum total of these decisions makes all the difference between greatness and mediocrity for the institution.

A second method of classifying decisions, particularly applicable in colleges and universities, is to separate them into those that involve the academic program and those that are non-academic. The distinction is based on the idea that, if decisions can be so classified, then the authority to make decisions can be assigned to the respective elements of the organizational structure. This classification system therefore becomes a basic point of reference in the assignment of institutional personnel to either the academic or the non-academic grouping. It is usually the most important element determining the structure of the official chart of administrative organization in a college or university.

The distinction under this method is probably more useful for classifying decision makers than for classifying the substance of decisions. It is very difficult to identify any large number of situations where the decision does not involve the academic program in some manner; and also situations that may seem purely academic in nature quite commonly involve such non-academic areas as finance, physical plant, and (so-called) non-academic personnel. For example, decisions about the salaries to be paid faculty members might be considered as purely academic, but certainly finance is involved and probably also public relations of a non-academic sort. Perhaps the most difficult problem presently and persistently facing many of the growing colleges and universities is automobile parking for faculty, students, and staff. Is this an academic or non-academic problem?

A third system of classifying decisions involves the separation into those that involve policy and those that involve action. Presumably action is merely the execution of policy, but sometimes action is taken first, with policy evolving from a series of more or less consistent actions. Or the necessity for a policy may be observed when there is a question about the consistency or appropriateness of an action.

Under this system of classifying decisions, the general rule is that policy decisions are made normally at a higher administrative level than decisions to take action. For example, the board of trustees of an endowed college or university may lay down certain policies to govern the investment of endowment funds. The officer or agency that actually manages the fund, however, makes decisions about the specific securities to be purchased (an action decision) in accordance with these policies. Or the entire faculty may make a policy decision about the qualifications of students that are to be admitted, and this policy may and probably should go to the board of trustees for official adoption. But the decision to admit a particular applicant is usually delegated to a subordinate in the administrative structure, the admissions officer.

In the research done in the early 1930's for the North Central Association of Colleges and Secondary Schools for the development of the new procedure in accrediting institutions of higher education, there was an extensive investigation into the location of decision-making functions in colleges and universities. The basic classification of the administrative situations was in terms of the making of policies and the taking of executive action in accord with the established policies. This classified list of administrative or decision-making situations proved to be a most illuminative introduction to the whole nature of the control and operation of a college or university. It was used for many years by examiners as one basis for judging the soundness of the administrative system in colleges and universities being reviewed for membership in the Association.

As in the case of classification on the basis of large and small decisions, classification on the basis of policy and action demands some tolerance of exceptions in the usual ideas about assignment of responsibility. A lower official in the administrative hierarchy may be confronted by a situation in which he has to make a decision without benefit of consultation, and for which he finds no policy. If he is of the stuff good administrators are made of, he will first consider the situation from the standpoint of policy.

A fourth method of classification for decisions, sometimes used in the literature of general administration, divides them into the routine and the critical. The terms programmed and unprogrammed may be used for essentially the same sort of distinction. A critical or an unprogrammed decision requires fresh thought and study, one for which adequate rules or precedents are not readily available. The routine or programmed is made by reference to established policies or precedents, requiring no re-thinking about the extended implications. The classification on this basis is not greatly different from that between large and small, or important and unimportant decisions. The classification on the basis of routine or critical, however, does have the advantage of focusing attention on the quality and quantity of the mental activity that must be involved on the part of the decision-maker.

While each of the four classification systems suggested is dichotomous, under none of the systems is the division a sharp one. Under any classification system decisions seem to be a continuum, ranging in a spectrum from small to large, or from routine to critical, from policy-making to action, and from academic to non-academic. There must therefore be considerable tolerance about any generalizations concerning the assignment of decision-making functions to personnel at various levels in the organizational structure of a college or university.

But who makes decisions concerning higher education in general? Every citizen, casting a vote for a Congressman or Senator, or for a member of the State Legislature, a Governor, or other elective officer, indirectly is making a decision that may affect higher education. Every young person who decides to go to college, and to a particular college, or who decides not to go to college, participates in a decision-making process that, in toto, has a most important impact on higher education. Every employer who sets academic requirements for newly employed personnel is affecting higher education. A striking instance of this last situation was the decision by the U. S. Armed Forces, shortly after the end of World War II, that commissioned officers must have the bachelor's degree as a minimum academic attainment.

More specifically, certain agencies outside the institutions of higher learning make decisions affecting colleges and universities. Four kinds of agencies may be noted.

1. Government officials, of all three branches of government--executive, legislative, and judicial--are among the most important of the outside agencies whose decisions affect higher education. They provide appropriations, often for specific purposes or with restrictions that the institutions must heed. They pass general laws that apply to colleges and universities, such as minimum wage scales or equal employment opportunities for women. Court decisions affect institutional policies in many ways, such as discrimination in admissions. Executive agencies require extensive statistical reports, and make decisions about grants of funds.

2. Philanthropic foundations provide grants of funds to some colleges, and deny grants to others. Some of the foundations encourage colleges to install new programs in the hope of quali-

fying for a grant. The foundations make studies and publish reports advocating changes in the direction of higher education, not without effect in many quarters.

3. *Media of public information*, perhaps not always consciously, tend to affect the operation of higher education. When the cartoonist of a widely read newspaper constantly caricatures the college professor as a wild-eyed nincompoop, clad in a disheveled academic gown and with his mortar board askew, the public image of college and university personnel is definitely affected. By contrast, when the Advertising Council of America sponsors a series of well done advertisements, paid for by their clients in industry and business, urging people to support the college of their choice, the impact is certainly significant.

4. *Churches* and other religious bodies have long been a highly important influence in the decisions about higher education in the United States. The first American colleges came into being through decisions by church groups, and this means of gestation and birth is still responsible for new colleges almost every year. In many colleges the church is still the ultimate authority over internal policies. Funds received through church groups are an important part of the financial support of many schools.

Inside the institution (or group of institutions in case the system of higher education is the operating unit) there are four fairly distinct groups that may have decision-making responsibilities: the governing board, most often known as the board of trustees or board of regents; the executive or administrative staff; the faculty members; and the students.

1. The *governing board* is commonly said to have a final authority over all the affairs of the institution. The specific powers of the board are usually set forth in the institution's charter or other legal authorization for existence. Property is usually held in the name of the board, and the board has the power as a corporation to sue and be sued in the courts.

Although the board legally has the final decision-making authority, in practice it inevitably has to delegate much of this authority to other officials or agencies within the institution. The extent of this necessary delegation is more or less directly proportionate to the size and complexity of the institution. In a very small college, for example, the board may consider carefully every person proposed for appointment to the faculty, spending considerable time on this matter before voting to approve recommendations laid before it by the president or dean. In a large university, most faculty appointments are given only a *pro forma* approval by the board for the sake of the record, or the authority to make final appointments to the lower academic ranks may be specifically delegated by the board to the president or other executive staff members.

Especially in state systems of higher education, where one board has charge of a number of colleges and universities, some of which may be quite large, the board typically has to rely on its own executive staff and on the staffs at the institution level for most of the decision-making process. The alternative is a set of prefabricated rules that impose a cramping uniformity on the institutions and that commonly lead to inefficiency and mediocrity. It is unfortunate, also, for the sheer magnitude of the operation to preclude attention by the board to the making of policies which should have a review at that level.

The device of using a board for decision-making at the highest level is the American way of safeguarding against executive tyranny or mismanagement. There is grave danger in risking too many such decisions to institutional executive officials, no matter how competent they may be, but this happens when a single board has a span of responsibility beyond that to which it can give adequate attention.

2. The *executive or administrative staff* of the college or university is usually the nerve center of the decision-making process at the local institutional level. Without doubt many decisions about institutional policies and actions originate with or are made by the executive staff. Almost all decisions of any consequence flow to it or through it.

In the past it has generally been held that the responsibilities of the governing board should be confined chiefly to matters affecting policy, while the administrative staff should be concerned primarily with executing policy. There are many instances in which a decision by a governing board on an action, that is, the application of a policy to a specific case, has been unwise.

This generalization about the division of decision-making responsibilities between the board and the executive staff is still valid. There are, however, many exceptions. Inevitably in a college or university many decisions on policy will be made by agencies or persons other than the governing board. It is only the "important" or "large" policies that are expected to come to the board for decision-making. But what is an "important" policy? And who makes the decision that the policy in question is sufficiently important to be brought to the attention of the board? Simultaneously there are situations in which an action-decision can be taken only by the governing board--for example, to hire a new president or to fire the incumbent president. But should the board take action on its own initiative to fire the football coach, or the dean of women, or an assistant professor of theology? "Certainly not," is the common answer. "The board should act only on a recommendation coming up through the executive staff."

In the popular mind, the administrative staff of a college or university is the chief or even the sole locus of the decision-making responsibility for institutional affairs. However, this is only partially correct, as revealed by a discussion of the faculty and students. The two groups next discussed constitute the chief limitations on the decision-making power of the executive staff within the institution.

3. The *faculty*, individually as members and collectively as a legislative body, constitutes the third group involved in the decision-making process within the institution. It seems almost axiomatic that decisions affecting the curriculum and the instructional process should be made by the faculty members, for it is in those areas that they are the experts. Furthermore, the *de facto* curriculum is only what happens in the classrooms of the professors. Also it may be agreed that the scholarly qualifications of a prospective appointee to the faculty can, under normal conditions, best be

evaluated by the faculty members in the discipline.

The professoriate has long held that it should be allotted responsibility for decision making not only in such recognized areas as curriculum and instruction, but also in other phases of institutional operation (such as budget determinations and new plant construction). On a national scale, the most vocal group representing faculty interests is the American Association of University Professors. This association has long demanded that the institutional faculty should have the right to elect its own representatives to the governing board of each college or university. But little progress toward this goal seems to have been made at the present.

Moreover I seem to detect a lessening of emphasis today on the "rights" of faculty members to participate more widely in the decision-making process on all matters affecting their institution. If I am correct, why should this be so? It might be taken as an indication that reasonable gains have been achieved, and that the faculty now does enjoy the decision-making powers that have been sought for it. Or more likely, the recent notable improvement in faculty salaries may have removed the most significant causes for complaint against the limited participation of faculty in decision-making. There is some interesting evidence on this point in the study completed a few years ago at New York University concerning faculty satisfactions and dissatisfactions. That study found that an unusually high level of satisfaction prevailed among faculty members in what were then the reasonably high salary brackets.

4. Students constitute the fourth group with more or less participation in decision-making in institutions of higher education. The standard pattern has been to allocate to the student group the decision-making authority in matters affecting only students (quiet hours in dormitories, fraternity and sorority rushing, and the distribution of receipts from activity fees). Yet even here much paternalistic influence or guidance from faculty and administrative staff has prevailed. But presently students are demanding participation in decision-making on a much wider range of institutional affairs then the traditional pattern provides. They want to make decisions about curriculum requirements, about retention and promotion of faculty members, about public lecturers brought to the campus, about the use of course grades to denote academic achievement, etc. However, we do not have an accurate assessment of the representativeness of the whole student body in such demands. Are these only the mouthings of a few agitators, or do they represent a broad movement, widely supported by most of the students? Already the use of the phrase "student power" (in an obvious analogy to "black power") raises some questions about the real origins and strength of the so-called student revolt.

It is perhaps an over-simplification of the situation, but I observe that the rise of the new demands of students for wider responsibilities in decision-making is coincidental with the loss of impact of a decision-making power formerly held by students--the power to decide what college to attend and to choose from a wide variety of institutions, all of which were eager to have him or her as a matriculant. In very recent years we have rather suddenly shifted from a buyer's market to a seller's market in college attendance. Or, more specifically, students heretofore could exert an important power over institutional authorities by refusing to attend one college and choosing another. That power has been diminished greatly; when almost every institution has plenty of good students, and when getting admitted to college, and to a particular college of the student's choice, involves facing a formidable barrier of selection devices. In the former buyer's market, institutions were particularly eager to attract students, and students were conscious of being wanted, a feeling closely akin to the feeling of power. Having lost this indirect power, students tend to seek more direct participation in the decision-making process.

Factors Affecting Decisions

What turns a decision one way or another in the decision-making process? The factors that affect decision-making may be grouped into four categories: personal factors within the decision-maker himself; pressures on the decision-maker from outside sources; precedents; and objective data or evidence.

The decision-maker needs to be aware of personal factors, such as his own biases and prejudices, which may sometimes lead to unwise decisions. Most subordinates quickly analyze their chief, learn about his biases, and anticipate how these may sway a given decision. In some cases the decision-maker's own self-interest in a situation may affect his decision. Higher education is remarkably free from instances in which an institutional executive has lined his own pockets because of decisions he has made, but such instances are not unknown. Much more subtle is the temptation to consult one's own ease and comfort, or that of close friends or relatives, when a decision might affect such interests.

Some decision-makers gamble. They tend to disregard all other factors and depend on their luck. Quite in contrast is the executive who likes to follow fairly fixed principles in making his decisions. With him also subordinates can rather quickly grasp the basic principles which he uses as guides, and they can successfully anticipate what kind of a decision he will make in most cases.

Any decision-maker comes under pressure from others regarding the determination of a policy or course of action. Associates and friends, many of whom he respects for the soundness of their advice, may bring pressure. More difficult is the pressure that can arise from those whose unfriendly action he fears and must guard against.

Decision-makers prefer to be considered consistent in their decisions. Therefore precedents bring a marked influence on the decision in almost any situation. Sometimes, however, precedents need to be overturned, in light of changing circumstances in a college or university. Oftentimes a review of the manner in which other similar institutions have met a situation provides the kind of precedent which a college or university can wisely follow.

One of the forms in which precedents may be cast is the institutional statement of its aims and objectives. Most colleges and universities have such a statement, carefully considered, officially adopted, and often printed on the early pages of the annual catalog. How much attention do decision-makers pay to this statement in arriving at their decisions? A cynic might reply, "Only when such a reference serves to bolster up an already conceived decision that may be subject to criticism by others." It is often quite convenient for the decision-maker to say, "But this is a liberal arts college, according to our statement of objectives, and therefore we must decide in this manner." One might also deduce, from the record of decisions made, another set of institutional objectives that might differ considerably from the printed list in the annual catalog. The whole question of the relation of institutional objectives to the decision-making process in higher education deserves more exploration than it has had, and certainly it is too complex to be treated here by more than a reference.

A fourth factor affecting decisions is objective evidence. An able executive will already be in possession of a considerable fund of information about any situation in which he is called upon to make a decision. But there is always the possibility that he may not have all the facts before him, and that some of his information may be unsound or out of date. This is where institutional research comes into the decision-making picture. I shall mention only briefly some of the kinds of objective evidence which the decision-maker needs.

The most common kind of evidence is probably statistical data relevant to the problem at hand. The data should relate to all pertinent aspects of the problem, and the analysis should be in sufficient depth to elucidate the areas that are of special significance for the decision to be made. Data on the present status of factors involved in the problem are most useful if they can be accompanied by reliable data of a normative sort, so that comparisons can be made with similar situations within the institution or in other institutions. In many cases it is quite helpful to have data in a time series, showing trends in various aspects of the situation of a period of months or years. Often it is useful to project past trends into the future, in the form of estimates of what the data will probably resemble at some period five or ten or more years ahead.

Another kind of information, which is often put into statistical form but which is somewhat different from the hard data obtained by counting dollars,consists of summaries of opinions. Such summaries, if pertinent to the problem at hand, accurately expressed, and carefully compiled, can be of great help in the decision-making process. It should be remembered that such summaries are merely a count of the varied decisions made by those included in the report. Summaries of opinion vary widely in quality. Many are only vague estimates. At the other extreme are polls obtained by carefully controlled random sampling, or that actually represent a total response from some defined universe.

Finally, the timing of decisions is a factor of utmost importance. To one who furnishes the raw material from which decisions may be made (for example, the director of institutional research in a university or college), one of the most frustrating experiences is to find that, after much pertinent information has been assembled, the decision-maker decides not to make a decision at this time. Now this, too, is a decision, even though it may result in no immediate change in the status quo and may relieve no tensions or permit no new planning.

Delaying a decision is a widely used tactic in administration. Sometimes the delay has the advantage, if the proper decision is expected to be unpalatable in some quarters, of keeping the opposition off balance. Or, a delay in decision-making may reflect one or more of several other possible situations. It may result from a belief that sufficient facts and evidences are not yet available to justify a decision. Or, the decision-maker may believe that the correct decision may meet more popular acceptance later. Quite commonly the decision-maker harbors a secret hope that the matter may solve itself without a decision. And frankly, some decisions are delayed because the decision-maker just cannot decide.

Most successful administrators have an uncanny sense about the timing of decisions. The right decision at the wrong time can wreck an excellent idea or program. A far-sighted executive must cultivate a great store of patience to wait for the proper time to announce a decision, and good ideas may be stored for months, years, even decades. Some weak executives support a good idea after others, more bold, have begun to accept or adopt it.

The right decision at the right time is the golden key to progress. It is a prime characteristic of good leadership. Presently, this characteristic is especially rewarding in the field of higher education. But even the leader who successfully times decisions must have advisers who suggest new lines of development without regard to their timing. This is the function of the prophet or the seer, which in modern parlance could be translated as the researcher.

28/29

THE ROLE OF INSTITUTIONAL RESEARCH IN DECISION MAKING

Thomas R. Mason
Director of Planning
The University of Rochester

John Dale Russell, in his concluding remarks, threw out the challenge to "the next speaker" to define the function of institutional research in the decision processes he had just described. He said:

> Even the leader who successfully times decisions needs to have about him advisers who can suggest new lines of development without regard to their timing. This is the function of the prophet or seer, to use two old and very respectable terms which in modern parlance could be translated as the researcher. Is this a proper function of institutional research?

As John Dale may have anticipated, my answer is an emphatic, yes! I will take the position that institutional research should be oriented to *planning*, another term that resembles the ancient functions of the prophet or seer.

Planning is the process of research, analysis, projection and evaluation of alternative courses of action that preceeds a decision. In modern organizations, where certain kinds of decisions involve significant costs and consequences when committed to action, planning has become increasingly formalized and specialized as a staff function.

The kinds of decisions that entail changes of policy, commitment of substantial resources, and growth of major consequence insist on deliberation and objective evaluation. The complex interrelationships among students, faculty, programs, facilities, and finance in the modern college or university exceed the comprehension of unaided judgment and intuition.

The so-called "managerial revolution" reflects the effort to extend the reach of judgment and intuitive understanding to predict the impact of change in one set of components upon others in the organizational system. Given added impetus by the rapid spread of the "Program Budgeting" movement through government and education, the notion of *comprehensive planning* in higher education is increasingly popular.

Looking to the computer as the instrument that will permit us to cope with the complexity of the institution, many institutional research people are trying to develop simulation models and "total" information systems. Although at the moment these new tools threaten to compound complexity rather than reduce it to manageable proportions for effective use in decision making, the effort will produce some useful results in the next few years.

The institutional research office should be equipped to assume a coordinating staff role in comprehensive planning. As the focal point of institution-wide data collection, and equipped with skills in the methods of research and analysis, institutional research presumably has built a solid foundation of information and understanding of the relationships between programs and resources. What may be missing is the *planning orientation*, the frame of mind that turns this knowledge and understanding toward the future and assesses the long-range problems and potentials of institutional development.

I should like to discuss some of the general features of the planning orientation, and then try to give some illustrations of how it may be brought to bear on the decision making and policy formation processes of higher education.

The Concept of Rational Decision Making

The concepts of decision making and planning are very closely related. March and Simon distinguish between "programmed" decisions--those for which procedures, routines, criteria, precedents, and rules are established so that the decision is almost automatic--and "unprogrammed" decisions, which require innovative solutions to new problems. (March and Simon, 1958.) They describe planning as unprogrammed decision making.

Modern man has a powerful drive to strengthen the rationality of decision making, especially in large organizations where order and predictability of behavior are essential to the viability and continuity of the organization. (This is also viewed as the process of bureaucratization.) The organizational man is constantly striving to program the decision process to gain predictability and control over the components of the organization. Nevertheless, the dynamics of change, both internal and external, continually require response to new conditions for which lucid decision criteria are not available. Planning is the attempt to minimize the uncertainty of change and the unpredictability of the future by formalizing the process of anticipation. In a sense, when the planning process is formalized, it is an attempt to program the process of unprogrammed decision making.

The planner strives to approximate the ideal of rational decision making. Meyerson and Banfield (1955) have summarized this ideal as follows:

1. the decision-maker considers all of the alternatives (courses of action) open to him; i.e., he considers what courses of action are possible within the conditions of the situation and in light of the ends he seeks to attain;

2. he identifies and evaluates all of the consequences which would follow from the adoption of each alternative; i.e., he predicts how the total situation would be changed by each course of action he might adopt; and

3. he selects that alternative the probable consequences of which would be preferable in terms of his most valued ends.

As organizations become more complex, the officers who bear the authority for making decisions require specialized staff to carry out analysis and evaluation of the consequences of alternative courses of action. The planning specialists must be sufficiently removed from day-to-day operational decisions to avoid what March and Simon (1958) call "Gresham's Law of Planning":

> Daily routine drives out planning....When an individual is faced both with highly programmed and highly unprogrammed tasks, the former tend to take precedence over the latter even in the absence of strong over-all time pressure.

On the other hand, if the planner is too insulated from the day-to-day flow of action in the institution, he will be out of touch with the realities that must be taken into account if planning is to be effectively useful in the decision process.

The cliche that "planning should be a continuing process" is meaningful when all of the participants in the institution's decision-making processes are possessed with the planning orientation--the frame of mind that constantly refers decisions and actions to a conception of what the institution is becoming, an image of the desired potential of the institution projected into the future.

The creation of this image and its continuing modification and elaboration is not the product of the planner alone; it is the product of the planning process in which all of the participants in the organization's decision making share. In a college or university, these participants are very numerous--the faculty, department chairmen, deans, administrative officers, trustees, state and federal officials, alumni leaders, and students.

In this kind of pluralistic organization, the achievement of a precise and coherent plan of what the institution ought to be, to which all parties will agree, is very unlikely. Clear and concise formulation of goals and objectives--one of the requisites of the ideal model of rational decision making and planning--cannot be completely achieved. (Banfield, 1962.) The effective planning images are more likely to be found in the minds of the institution's leaders rather than in the documents called "master plans" that we find collecting dust on the shelves of so many presidents' offices.

An effective planning process, however, needs a conceptual system upon which the bits and pieces of consensus can be hung and against which the flow of decisions that have enduring consequences may be tried and tested. The institutional research office is a good agency to devise and maintain this planning conceptual system. All too often, however, the IR staff is so consumed with the burdens of discovering and evaluating what did happen, there is no time to project and assess the course of past and current events into the immediate future, let alone the longer-range. For institutional research to assume the comprehensive planning role, it must be injected with a strong dose of the prophetic urge. As I will try to show, this powerful drug may have some undesirable side effects of which you should be aware before you become addicted.

The Uses of Prophecy

The term "prophecy"--suggested by John Dale Russell as translatable into research--hints at something of the mystical prescience of the ancient wise man or medicine man. The notion transcends the hard facts we ordinarily seek in our research. This transcendence is exactly what is required in the planning orientation, an ability to go beyond the certified data about the past and present to envision the possibilities of the future.

Prophecy of the course of future development requires a feel for the intangible goals, values, prejudices, and preferences of the actors involved in a cooperative social enterprise, such as a university or college. The successful prophet perceives the tendencies of this web of values and goals and shapes his prophecies to forecast their fulfillment.

A well-conceived image of the future that takes into account the nature of the prevailing values may become a kind of self-fulfilling prophecy because for most of the participants in the organization, the more distant future appears highly uncertain. If a projection of some set of factors generally coincides with their value expectations, it is viewed as a valid statement about the future, reducing some of the uncertainty. Literally thousands of major and minor decisions may be made by individuals in the organization in anticipation of that projected outcome. As a result of the compounding of these decisions, aimed toward the outcome projected, the probability of its fulfillment is greatly increased.

If, however, a prophecy does not please the king, the wiseman's head may roll. If a projection does not forecast the outcome preferred by a significant body of influential participants in the organizational system, it becomes a matter of controversy. This controversy is one of the potential side-effects of planning. Some controversy and conflict can be accommodated in a well-ordered society which has mechanisms (e.g., committees) for the resolution of conflict. This kind of conflict may even help to clarify institutional goals and values upon which there is a reasonable consensus. Other kinds of conflict involve value cleavages so fundamental that the controversy may be destructive to the cooperative system. Every skillful administrator seeks to minimize destructive conflict, and the timing of decisions mentioned by John Dale Russell is an essential skill in tempering the impact of conflict on the organization.

Planning, then, involves more than simply projecting the past into the future. It cannot be limited by predictability, as the prediction of natural phenomena is viewed by the scientist. Planning is deeply involved with the complexity of human values and interests, and we all know how unpredictable these can be. The planner must be aware of the subtle processes of goal formation and skilled in anticipating the value-responses of the participants in the organizational system.

This is not to say that the work of the planner cannot be objective; objectivity of evaluation, a facing of the consequences, good or bad, is a vital attribute of the staff work supporting the planning process. This emphasizes one of the other hazardous side-effects of the prophetic urge. Active participation in the decision and policy formation process implies commitment to an objective or a course of action once the decision is given the stamp of authority. Once commitment is made and must be justified and defended, the danger of bias of data, mentioned by Dan Robinson, is present. This may lead to loss of the attribute of objectivity we all seek to have ascribed to our research. Some IR people seek to insulate their operations from this danger by trying to remain detached from active involvement in decisions: "We just state the facts, and let the policy-makers decide what to do with them."

I think this position is artificial when institutional research is expected to serve as the handmaiden of administrative policy. Uninterpreted quantitative facts are frequently useless, often dangerous, and always subject to misuse standing by themselves, simply because they are incomplete. Objectivity requires the _full_ evaluation of the problem, including the identification and assessment of the intangible, qualitative values (benefits or utilities) that must be weighed judgmatically against the more measurable factors, such as cost. Qualitative values are facts and although they are more difficult to identify and communicate, they must be tied to quantitative analysis before it is objectively complete.

The objectivity of research applied to the planning process depends on the recognition that the role of institutional research in decision making is a _staff service function_. Objectivity is a quality of the service. The service is provided to the president, his principal executive officers, the deans, department chairmen, and faculty committees who are, in various combinations in various situations, the generators of institutional policy and plans. Academic and administrative leadership is the consumer of organized information, analysis, and evaluation produced by staff specialists (whatever they may be called) who conduct research, integrate information, and evaluate the costs, values, and consequences of alternatives on the agenda of policy formation.

The degree and significance of the role of institutional research in decision making probably depends primarily on the administrative style, outlook, and expectations of the president and his immediate executive officers. The living processes of decision making operate in a complex and fluid system of personality, situation, and environment unique to each institution and changing with changes in personnel in the key academic and administrative roles. If these officers do not care to have you involved in the game, you must stand on the sidelines. To get into the game, you may have to prove your worth by being waterboy or cleaning up the stands. You can be prepared, however, when the big moment comes, and the coach tells you to suit up because he needs help.

It takes time for a new institutional research office, or for an old hand in a new position, to build a solid basis of information and understanding about the institution. I have found that this process moves much more quickly if one begins with some conceptual framework around which one can start a coordinated flow of information and a set of analytic studies that describe the interrelationships between the components of the institution. Since the institutional research function should involve other staff, administrative officers, and faculty, the conceptual framework can help to clarify their role in the research effort and strengthen their contributions.

As an illustration (rather than a model) of a conceptual framework for research and planning, I will offer the schemes that I have recently been trying at the University of Rochester.

A Conceptual Framework for Planning

In a series of charts at the end of this paper, I have attempted to diagram the linkages between programs and resources that we are trying to comprehend in a systematic research and planning framework. They are intended to be suggestive, and I will assume that they are essentially self-explanatory.

The essential value of this kind of scheme is that it helps to clarify the interrelationships among the major components of the institutional system. As we build information and work on planning or policy problems related to one component, we try to assess its effects on the others. In the process, we are building our understanding and information about the other components and the nature of the relationships. This not only helps to trace the consequences of one kind of decision on other parts of the system as part of the analysis and evaluation of a decision problem, it also builds a store of knowledge to improve our projections and estimates as we proceed.

For example, our Office of Planning and Institutional Studies is responsible for coordinating the campus development plan and the facilities development program. These tasks place a heavy demand on our time that could easily diminish our research effort and overall planning perspective. Facilities planning is a good basis for organizing a comprehensive program planning effort because the costly and enduring commitments involved in capital facilities demand long-range thinking.

One of the first tasks of the office when it was organized in September, 1964, was to coordinate the preparation of a long-range campus development plan. The University had already established a fairly firm development program to 1975 as the basis for a major capital fund raising campaign. Its general development goals, in terms of student enrollments, faculty development, and broad program content, were pretty well determined, and we could start several leaps ahead.

The campus development planning process serves to bring out for discussion and evaluation a host of issues and values regarding the character of the institution, its organizational patterns, and the functional relationships among the departments, colleges, and general facilities. The need to secure immediate decisions on the sites of a number of new facilities--a very enduring kind of decision--makes the planning process meaningful to the participants and not just an exercise in drawing pretty pictures. It is a process that raises the most fundamental questions about the scale, content, and organization of the institution. It should be used as an integral part of the institutional self-study and not left to the superintendent of physical plant and the consulting architects or planners.

The first major facility on the priority list was the expansion of the main library. The programming of this facility was organized as a major series of decision problems: what kind of library should it be? The question requires discourse on the kind of University it would serve. Hard decisions are required in programming a facility. How many reader stations? This University has a very highly selected student body. Its four-course system assumes extra time spent in library work. About half of its graduate students and faculty will be using the library as the instrument of their research. Answer: we will provide reader stations for forty per cent of the full-time undergraduates, fifty per cent of the graduate students using this library, etc. How many volumes? This is a major research institution. Double the capacity by 1975. What must be the materials acquisition budget to accomplish this over the next ten years? What are the acquisition and cataloging staff requirements and costs? What other personnel will be required? What changes in library technology must be anticipated?

The program and operating ramifications of this process are almost infinite. The problem could have been studied for years without reaching conclusions, but the need for decision was there. And with a logical and coherent format for getting program decisions agreed upon, the programming was completed in about three months. The decision process, of course, must continue through design, since all kinds of problems of organization and function come up that require evaluation. Applications for federal grants and the continuing problems of cost control required further research and evaluation.

By early 1965, the campus development planning and library programming activities had produced an amazing amount of information and, more important, understanding of the images the university officers and faculty had in mind of what the university should become. In the process, it had become clear that much detail remained to be processed to continue the programming and operational planning phases of the scheme (Figure A).

We wanted to produce complete ten year space projections for all departments to serve as the basis for programming additional facilities and the assignment of space in existing facilities. The first step was to initiate a major campus-wide research project we called the Instructional Program Planning Study. The academic departments had previously done a considerable amount of work on planning their faculty development in terms of estimating the minimum number of faculty required of their programs, with sufficient breadth and depth to comprehend the special disciplines they expected to cover. The Instructional Program Planning Study asked them to project the development of their course offerings in light of their faculty development programs (Figure B). They were asked to anticipate the structure of the courses, in terms of type of instruction, contact hours, and maximum desired class sizes. The evaluation, analysis, and modification of the departmental responses provided us with a rich store of information about the composition and structure of the curriculum. Supplemented with studies of current course loads and structures, these data were used to develop the parameters for a simulation model of the instructional program, which has become a key element in our planning effort (Figure C).

The programming of space requirements has served to evaluate and project the research programs, staffing requirements, and special characteristics of the academic departments. These evaluations and projections serve as input for the conceptual framework, contributing a broadening comprehension of the dynamics of the institution--and revealing areas of informational weaknesses to be placed on the agenda for study (which, by the way, is growing rather long).

In an undergraduate housing study, a range of choices from single-room, six-man suites (the high cost alternative) to double-room, gang-style dormitories (lower cost) was brought before randomly selected groups of students in an experimental decision game situation. The students evaluations, recorded by a graduate student observer, played a major part in the decision, along with the discussions it evoked among student service personnel, business office personnel, and administrative officers. The higher cost alternative was chosen, deliberately, on the basis of the values gained.

Other facilities-based studies on graduate housing, special research facilities, a student center, a major science building, and traffic circulation, have been organized as decision problems evoking the appraisal of program values against costs.

This year the accumulation of information and evaluation that has been gained by this process was used in the preparation of a new ten-year financial plan. The original development targets were due for reappraisal, and the projections of operating costs needed updating. It was decided that a series of 1975 enrollment sizes and mixes should be tested in terms of operating budget revenues and expenditures from 1967 to 1976. Using nine combinations of undergraduate/graduate enrollment sizes and mixes, we put these through a generalized simulation model (illustrated in Figure C) for the colleges of arts and science, engineering, business, and education. The output of faculty requirements was used as the starting point for the build-up of the education and general income and expense projections. Manipulation of the parametric variables, such as class size and average faculty teaching load, a kind of sensitivity analysis, served to evaluate the impact of these critical factors on faculty requirements and hence cost. This exercise has uncovered a whole series of policy questions that need further study in depth. The use of graduate teaching assistants, for example, has been the subject of more intensive scrutiny. The size of the full-time undergraduate enrollments in the professional colleges relative to program offerings and cost were pinpointed as needing review by the college deans and their faculties. Thus the long-range problems of fund-raising, tuition policy, and student aid are to receive full scale evaluation in the coming year. The ramifications of this kind of planning study even surprised me, and the value of a systematic planning framework was fully demonstrated.

The Range of Participation

This kind of planning and research operation is capable of evoking the participation of faculty, students, and administrative staff as well as the executive officers. The decision evaluation process that we try to apply to the problems we tackle leads us directly to the people who can give us information, advice, and an expression of values. We do not hold referenda, convene the faculty senate, or hold a general assembly of the students to discuss these issues; we do not even have a faculty planning committee. When we cannot anticipate a response, we consult. We have managed to avoid serious conflict so far by anticipating it and seeking to alleviate it through discussion and negotiation. (Perhaps we just have not hit the hell button yet.) We find ourselves acting as intermediaries between the faculty and the administration. We may find ourselves pleading for an increase of budget to provide for some faculty desire, then turn to the effort to persuade a department chairman to accept a constraint or limitation.

By engaging the faculty in the process of long-range program planning, in fact, we are putting the weapons of analysis into their hands with the hope and expectation that they will do a better job of justifying their budgetary and space requirements and a better job of departmental planning. We hope that in time the systematic approach to planning will infect the entire academic and administrative organization. If this occurs, it may dispense with the need for an "academic civil service," such as that called for by Rourke and Brooks (1966), to counter the weight of bureaucracy.

A very wise university administrator once told me that every administrator should try to work himself out of a job. Although I do not really expect to achieve it because the coordinating staff role in institutional research planning probably is here to stay, my ambition is to find ways of so diffusing the planning orientation into the entire organization of the university that the Director of Planning and Institutional Studies can concentrate on being a prophet and seer.

REFERENCES

Banfield, Edward C., "Ends and Means in Planning," eds. Sidney Mailick and Edward H. Van Ness, *Concepts and Issues in Administrative Behavior*, 1962.

Barnard, Chester I. *The Functions of the Executive*, 1938.

Branch, Melville C. *Planning: Aspects and Applications*, 1966.

Bross, Irwin D. J. *Design for Decision: An Introduction to Statistical Decision-Making*, 1953.

Corson, John J. *Governance of College and Universities*, 1960.

Easton, David. *A Systems Analysis of Political Life*, 1965.

March, James G., and Herbert A. Simon. *Organizations*, 1958.

Meyerson, Martin and Edward C. Banfield. *Politics, Planning, and the Public Interest*, 1955.

Mason, Thomas R., "An Inverse Relationship: The Uses of Facilities Planning for Institutional Research," *Research on Academic Input*, 1966.

Mason, Thomas R., "The Dynamics of Planning in Higher Education," *Proceedings of the Eastern Association of College and University Business Officers* (December, 1966).

Novick, David, ed. *Program Budgeting: Program Analysis and the Federal Government*, 1965.

Riesman, David, "Planning in Higher Education: Some Notes on Patterns and Problems," *Human Organization*, XVIII (1959), 12-16.

Rourke, Francis E., and Glenn E. Brooks. *The Managerial Revolution in Higher Education*, 1966.

Simon, Herbert A. *Administrative Behavior: A Study of Decision-Making Processes in Administration Organization*, 1957.

Wallis, W. Allen, and Harry V. Roberts. *Statistics: A New Approach*, 1956.

Wildavsky, Aaron, "The Political Economy of Efficiency: Cost-Benefit Analysis, Systems Analysis, and Program Budgeting," *Public Administration Review* (Dec. 1966), 292-392.

Williams, Harry. *Planning for Effective Resource Allocation in Universities*, 1966.

T.R. Mason
November 1966

Figure A

The Basic Linkages of Programming & Planning in Higher Education

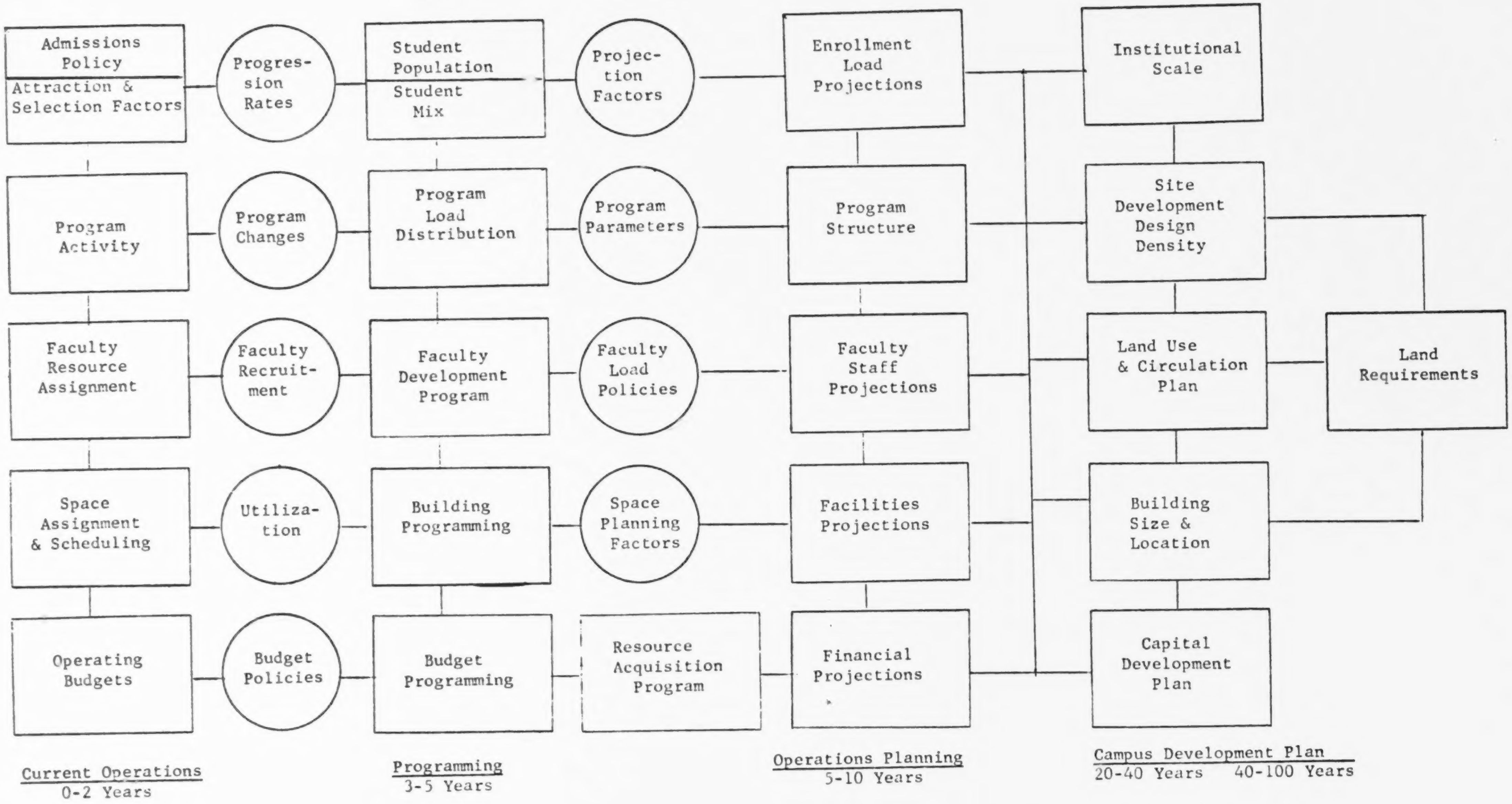
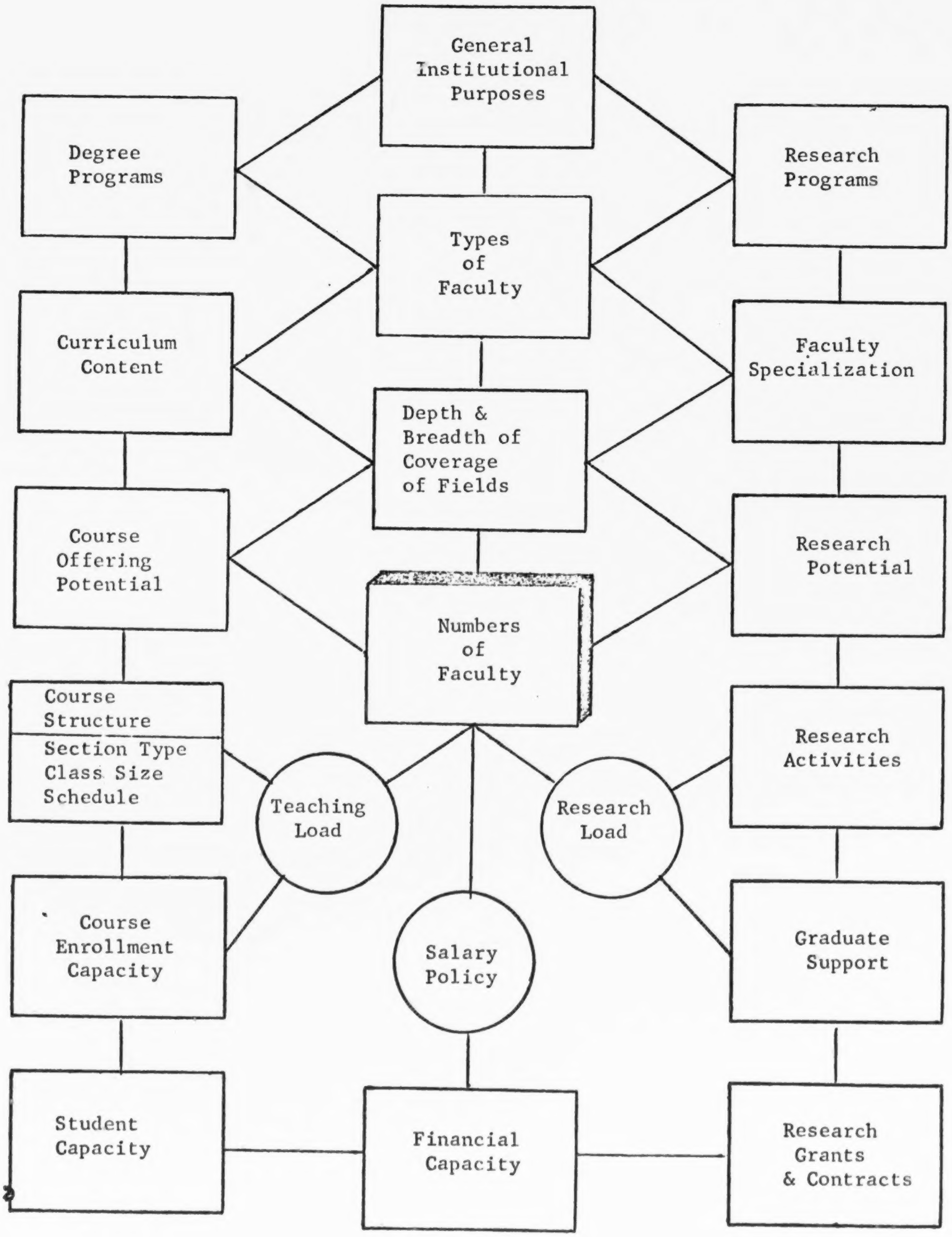

Admissions Policy
Attraction & Selection Factors
Progres- sion Rates
Student Population
Student Mix
Projec- tion Factors
Enrollment Load Projections
Institutional Scale
Program Activity
Program Changes
Program Load Distribution
Program Parameters
Program Structure
Site Development Design Density
Faculty Resource Assignment
Faculty Recruit- ment
Faculty Development Program
Faculty Load Policies
Faculty Staff Projections
Land Use & Circulation Plan
Land Requirements
Space Assignment & Scheduling
Utiliza- tion
Building Programming
Space Planning Factors
Facilities Projections
Building Size & Location
Operating Budgets
Budget Policies
Budget Programming
Resource Acquisition Program
Financial Projections
Capital Development Plan
Current Operations
0-2 Years
Programming
3-5 Years
Operations Planning
5-10 Years
Campus Development Plan
20-40 Years
40-100 Years

Figure B

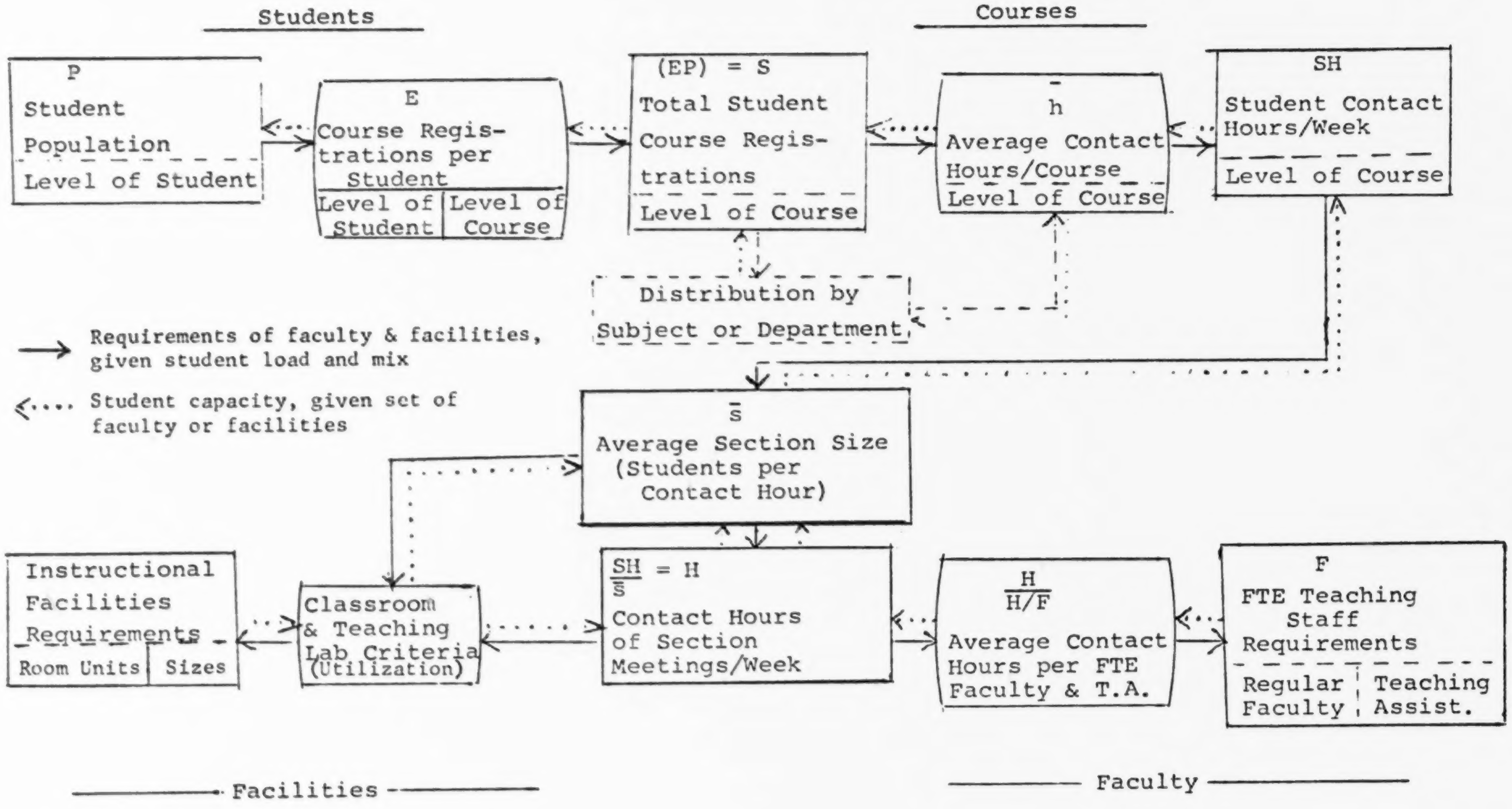
Program Structure & Faculty Resources


General Institutional Purposes
Degree Programs
Research Programs
Types of Faculty
Curriculum Content
Faculty Specialization
Depth & Breadth of Coverage of Fields
Course Offering Potential
Research Potential
Numbers of Faculty
Course Structure
Section Type
Class Size
Schedule
Research Activities
Teaching Load
Research Load
Course Enrollment Capacity
Salary Policy
Graduate Support
Student Capacity
Financial Capacity
Research Grants & Contracts

T.R. Mason
November 1966

Figure C

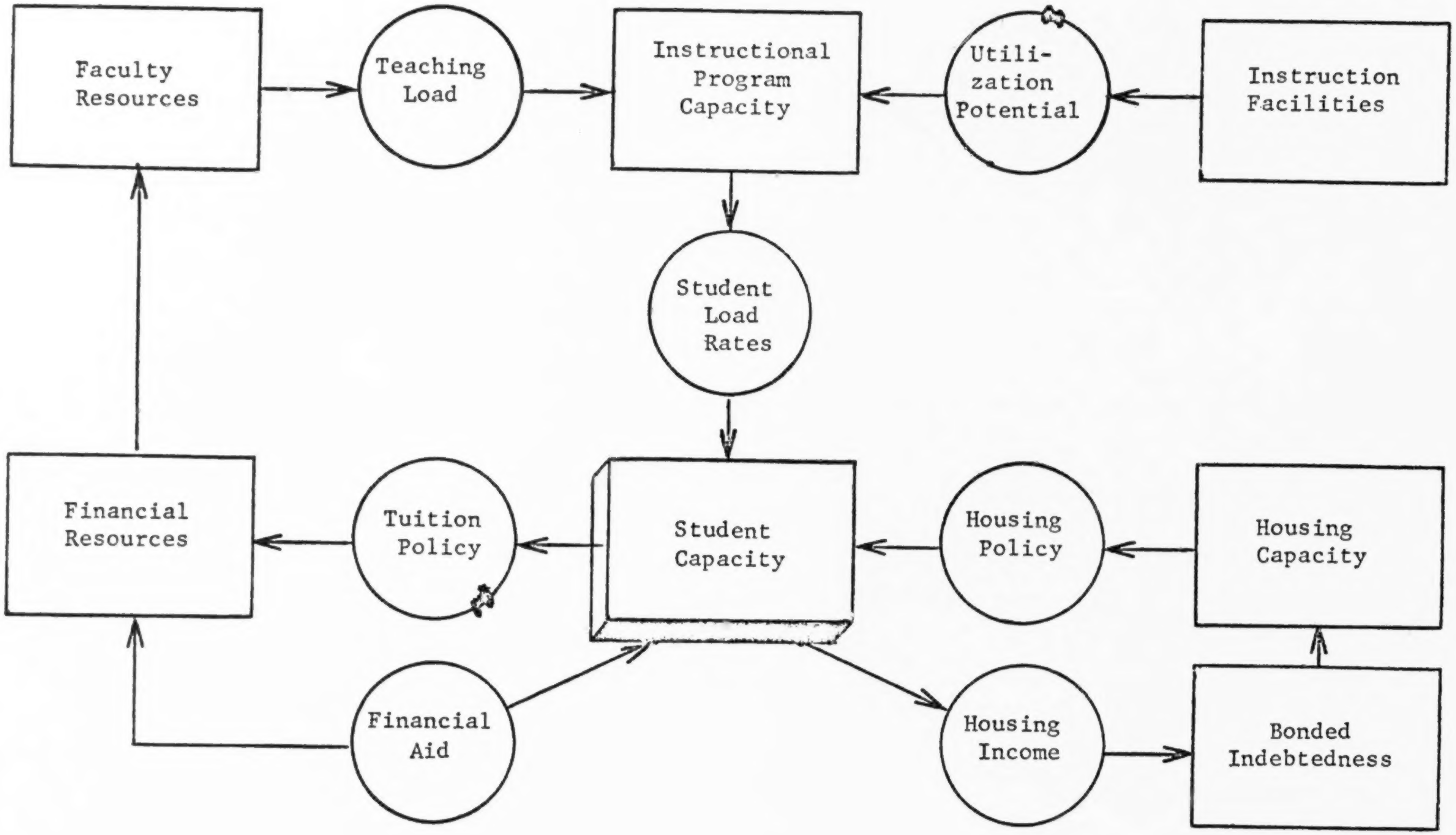
Flow Chart: Model of University Instructional Program


Students
Courses
P
Student
Population
Level of Student
E
Course Registrations per Student
Level of Student
Level of Course
(EP) = S
Total Student Course Registrations
Level of Course
$\bar{h}$
Average Contact Hours/Course
Level of Course
SH
Student Contact Hours/Week
Level of Course
Distribution by Subject or Department
Requirements of faculty & facilities, given student load and mix
Student capacity, given set of faculty or facilities
$\bar{s}$
Average Section Size (Students per Contact Hour)
Instructional Facilities Requirements
Room Units
Sizes
Classroom & Teaching Lab Criteria (Utilization)
$\frac{SH}{\bar{s}} = H$
Contact Hours of Section Meetings/Week
$\frac{H}{H/F}$
Average Contact Hours per FTE Faculty & T.A.
F
FTE Teaching Staff Requirements
Regular Faculty
Teaching Assist.
Facilities
Faculty

University of Rochester
Planning
1-66

Figure D

The Dynamics of Student Capacity


Faculty Resources
Teaching Load
Instructional Program Capacity
Utili- zation Potential
Instruction Facilities
Student Load Rates
Financial Resources
Tuition Policy
Student Capacity
Housing Policy
Housing Capacity
Financial Aid
Housing Income
Bonded Indebtedness

T.R. Mason
November 1966

INSTITUTIONAL RESEARCH AND THE INSTRUCTIONAL PROCESS

Stanley O. Ikenberry
Dean, College of Human Resources and Education
West Virginia University

The topic, "Institutional Research and the Instructional Process," requires certain assumptions about institutional research and institutional researchers as well as reasonable definitions of the instructional process. There are many definitions of institutional research and institutional researchers but variation along one major continuum tends to override the others. John Stecklein referred to the distinction in his presidential address of 1966 by quoting Henry S. Dyer's contrast of the need for theoretically-oriented, long-term research, free from the daily demands for immediate problem solving, and the contrasting vision of institutional research as an arm of the president or executive vice-president's office, with the primary goal supporting decisions on the use of financial and other scarce institutional resources.[1]

Sam Baskin betrayed his position in an address before this association in 1964 in which he argued that institutional research should be considered a sub-specialty of educational research.[2] Lou Mayhew last year suggested that institutional researchers were the captives of hard data and that when one considered "what is really known about the central structures of higher education the names which come most quickly to mind are not... devoting their full professional talents to institutional research...."[3]

Much of the thrust of this paper is related to this fundamental debate on the nature and purpose of institutional research and institutional researchers.

Let us turn for a moment to the definition of the "instructional process." To single out the instructional process for consideration would suggest that there are other primary processes or functions carried on by universities beyond which we call instructional. Such a point of view could be defended in terms of our functional nomenclature of instruction, research and public service. But if one views the function of a university from a broader perspective, it is possible to suggest that the instructional process is, indeed, the function of a college or university. Some 115 years ago, John H. Newman set forth his classical definition of a university, stating:

> A University training is the great ordinary means to a great but ordinary end; it aims at raising the intellectual tone of society, at cultivating the public mind, at purifying the national taste, at supplying true principles to popular enthusiasm and fixed aims to popular aspiration, at giving enlargement and sobriety to the ideas of the age, at facilitating the exercise of political power, and refining the intercourse of private life. It is the education which gives a man a clear conscious view of his own opinions and judgments, a truth in developing them, an eloquence in expressing them, and a force in urging them.[4]

Over a hundred years later Nathan M. Pusey suggested that:

> Our job is to educate free, independent, and vigorous minds capable of analyzing events, of exercising judgment, of distinguishing facts from propaganda and truth from half-truths and lies, and--in the most creative of them at least--of apprehending further reaches of truth. It is also our responsibility to see that these minds are embedded in total persons who will stand with faith and courage, and always, too, in thoughtful concern for others. We must all of us, at every level in education, work together to do this job.[5]

Neither Newman nor Pusey distinguish between instruction, research, and public service, for the functions are inseparable. Although directed toward a distinctive clientele, using different methods and conducted in a different context, our programs of instruction, research, and public service are each designed to discover, synthesize, share, and preserve human understanding. Thus, I define the instructional process in the broadest of terms to include the whole of institutional purpose.

Following from these positions, the argument may be synthesized by suggesting that institutional research is the study of the instructional process; that is, its purpose is to study the institution and its mission. Institutional researchers engaged in research and development activities which could not be classified as relevant to the instructional process (including research and service) may be engaged in essential functions, but are not engaged in the conduct of institutional research. In short, institutional research at its best is the analysis of the instructional process.

Turning now to a second phase of analysis, the contributions of institutional research over the last few years may be examined with the aid of a system of major categories which may suggest certain areas in which institutional research has made substantial contributions and other areas in which the long-term promise for American higher education may be the greatest, but the past contributions the most scant. The suggested major categories include: (1) analyses of input, (2) analyses of process, (3) analyses of context, and (4) analyses of output.

Beginning with an examination of analyses of institutional input, it is clearly in the area of input analysis that institutional research can boast of its major accomplishments. It is with input analyses that institutional researchers have been most closely identified. We can boast solid studies of student characteristics, faculty characteristics, space and facilities analyses, finance, and the curriculum. Each of these areas represents a major input source in American higher education.

In the area of student characteristics, for example, the contributions of Nevitt Sanford, T. R. McConnell, Paul Heist and others have been substantial. The work of Dorothy Knoll, Warren Willingham, Bill Warrington, Irvin Lehmann, and our own Jim Montgomery deserve mention. In fact, most offices of institutional research with a longevity of two or three years could likely produce a reasonably well done study of student characteristics at the local level and could suggest how the study had been helpful in influencing sound educational change.

In the area of faculty characteristics, the work of the Minnesota group led by John Stecklein is illustrative and well known. Studies of faculty characteristics including degree status, salaries, distribution among ranks, size, source, and other factors have been carried on in great numbers across the country.

Space and facilities analyses are also commonplace thanks to some of the initial work carried on by Russell and Doi, later contributions by Bill Fuller and Harold Danhke, and others who have continued to refine the methodology. Work in the area of financial analyses such as that of John Dale Russell some years ago in New Mexico and Doi in Colorado, contributions of James Miller, John Millett, and the Swanson-Still study--these and other efforts have given us a new and sophisticated expertise in the area of financial analysis.

Even the input of curriculum has come under inspection, stimulated by the work of Ruml, Dressel and McGrath, but now joined by others. If one were to walk on to most college and university campuses today and ask which office within the complex structure might be expected to carry out studies of student characteristics, faculty characteristics, studies of space utilization and facilities planning, financial analyses, or curriculum, he would probably be directed to the office of institutional research. Input analyses such as these conducted by offices of institutional research or those outside of them will continue to be important.

Analysis of student input, faculty input, the input of space, financial input and curriculum input, however, and the complex interaction of these input variables in studies of class size, faculty load and instructional cost, represent an accomplishment of an earlier day and no longer approximate the professional frontier. Although we must continue to accord close scrutiny to these input variables, we now run the risk of suffering at the hands of our own success in which the image of who an institutional researcher is and what he does is restricted to the endless replication of input studies, on the one hand, and ad hoc management studies on the other.

If the range of potential contributions becomes narrowed unduly, and the image of what institutional research is and does becomes solidified, American higher education will find it necessary to create still another new mechanism to accomplish the broader research and development functions toward which we will be moving.

What of the so-called instructional process? Suggested below are four process areas in which we have made modest contributions in the past, but which, at the moment, remain somewhat protected from the attention and resources of offices of institutional research. Although others could be mentioned, of special concern might be:

1. An examination of the effectiveness of various remedial or other special purpose instructional programs;
2. A functional analysis of supplementary services such as student counseling centers, residence halls systems, etc.;
3. Greater attention to the instructional processes in the classroom itself; and
4. Increased development activity in an area that is being called educational engineering.

Looking briefly at each of these areas, the literature in general, and the reports of offices of institutional research in particular, offer little hard evidence with reference to our various attempts to improve the achievement of sub-standard students.

Although one can point to isolated examples, the mandate of our offices, our self-concept, or our competencies and resources do not appear to push us in those directions. West Virginia University recently completed a study testing the effects of the common treatment of a reduced academic load and a special "how to study - reading improvement" course in the improvement of academic achievement and in the reduction of withdrawal rates of failure-prone freshmen. Our findings suggested, for example, that when applied in the absence of other measures, a reduced academic load may have negative effects on student performance and actually accelerate the rate of withdrawal.[6] We were able to find only one other directly relevant study in the literature.[7]

While I am sure one can cite examples of studies designed to evaluate the effectiveness of various remedial measures to improve student achievement and performance, most of our programs in these areas are supported by conventional wisdom and have not been subjected to careful scrutiny.

You may have read in the April, 1966 issue of the Review of Educational Research[8] a concluding statement of the state of student personnel programs in higher education. It suggests that basic deficiencies existed in the quantity and quality of research in this area, and points out that staff members who are most interested in student affairs programs are likely to be practitioners who have neither the time nor the training to plan and execute complex research. Moreover, it concludes that administrative and financial support and the collaboration of social scientists (and I would add institutional researchers) are essential to fill the apparent gap.

I would also join with Sam Baskin in suggesting that we must go into the college classroom much more than has been the case.[9] McKeachie summarizes well the state of research on the methods of teaching, comparing the effectiveness of the lecture method versus the discussion method, the influence of class size on the quality of instruction, and the inter-action between selected student and teacher characteristics and academic achievement.[10] But we have not gone very far beyond this obvious deadend. Perhaps we should be observing behavior using a system such as Flanders' inter-action analysis. Or, perhaps we should be exploring the implications of techniques such as micro-teaching and modeling as developed at Stanford and elsewhere. Unfortunately, I see little of it.

Turning to still a third area of the so-called instructional process, we should be more involved in the so-called educational engineering or systems development. We have a very unfortunate and parochial view of learning which is reflected in the nature of the research which we conduct on the instructional process. We begin with assumptions about a certain number of fixed seats in a classroom, a blackboard, a few audio-visual aids, a podium for the lecturer, and up-to-date textbook and a fifty minute period; then we begin to look within that rigid framework for possible variations. In fact, we have already destroyed 95 per cent of the potential variance.

Our orientation should be much more flexible and should search for the optimum educational environment with a minimum of restrictions assumed. The fifty minute period, the textbook, certainly the lecturer, and probably the class and the classroom itself should not be simply assumed as given; but should be varied, modified and challenged. Earlier work at Miami University, Pennsylvania State University, and others represent the directions towards which we may be moving in the use of instructional television, the studies in the use of programmed instruction and computer assisted instruction, the work done in the area of simulation systems, learning laboratories, and other approaches to teaching and learning.

We can expect in the foreseeable future a shift in educational investment. The educational institutions from kindergarten through the graduate schools will make much larger investments in complex instructional systems (both hardware and software), with greater emphasis on the potential for reproducibility, stimulus control, sequence modification, and a response record. Gradual improvement through a continuous cycle of research and development, characterized by some as the educational engineering approach, will be on the increase. In short, educational research and consequently institutional research will need to devote more attention to what happens to the learner and to the design and implementation of learning systems, with perhaps less attention to the way in which an individual instructor practices his art.

Whether on the traditional or emerging model, however, there is little good research on the instructional process and, with few exceptions, the contributions have been made by those outside of the institutional research community. To allow such a deficiency to continue must suggest the eventual decline of offices of institutional research and institutional researchers in the higher education scheme.

There is a third major category of analysis appropriate for offices of institutional research which I would describe as analyses of "context." Colleges and universities make certain decisions regarding inputs, they sustain processes, and they operate in various contexts, including location, structure, governance, size, etc.

Reviewing the literature a few years ago, I found a few papers published by the Western Interstate Commission on Higher Education as a result of their conference on "Academic Administration" and an issue of the Harvard Educational Review, with articles by Neil Gross and Ross L. Monney which considered the complex issues of university structure and organization. Mainly, however, we have failed to go beyond the talking stage in our analysis of institutional structure in American higher education. For example, as almost every college and university in the country experiences growth and as changes in institutional function occur, we continue to be befuddled by a lack of understanding of the relationship of these variables to institutional structure, organization, and administration. The current move to invest millions of dollars in the establishment of experimental colleges within the framework of larger universities and the implementation of other organizational adjustments is made without any clear understanding of the problem, let alone the solution. It is singularly appropriate that offices of institutional research accept major responsibility for deriving such data and suggesting such conclusions as perhaps possible.

But, not all is a void in the area of context analyses. Among the obvious positive examples in this area is the work by Pace and Stern and others who have examined the environmental press or institutional climate for learning of colleges and universities. As a result, higher education views with much more sophistication the nature of the academic community and the major variables influencing the environment for learning.

Yet, although we have spent considerable time and effort on organizational studies within single state systems, very little institutional research effort has been directed to comparative analyses of contrasting systems of organization, and there are few case-study analyses of distinctive state systems of higher education organized under atypical models. Colleges and universities operate within a governance context. A small private liberal arts college may report to a single board with single responsibility for that college. On the other extreme, an institution of higher learning may find itself as one among a growing complex in a total state system. Here again, if offices of institutional research are to be of value in the coming decades, they will need to take a broader view of the problems associated with the super structure and governance of institutions of higher learning.

As a final example of needed analyses in the area of context, Hubbard points up the need to understand more clearly the influence of the geographic setting. Like other aspects of our society, colleges and universities are moving to the city. The trend is so obvious and so irreversible that we may neglect its importance and significance. I suspect several important differences in context exist between higher education as carried out in 1920 in a basically rural setting, and higher education as carried out in 1970 in an increasingly urban and suburban setting.

Hence there are major gaps in our understanding of the context in which our institutions operate. Organizational structure, the impact of changes in size and function, the influence of the geographic setting, the nature of the environment for learning, the various systems of governance, these topics ought to claim a much larger share of our time.

Considering briefly the fourth and final aspect of our model, if the quantity and the quality of our analyses of institutional output matched that of our input analyses, higher education would have entered a new era. As a practical matter, however, our studies of output are few. Too frequently we settle for measures of intermediate output criteria such as student credit hours or the number of students taught, or the number of speeches given, or the number of articles published. We are short on the more substantive analyses much closer to the ultimate purposes of colleges and universities.

It is true that we have had more than our share of studies of student withdrawal, and, although negative, it is one illustration of institutional output. We have used several intermediate measures of student achievement, usually related to a single examination or perhaps a single course, or at best a cumulative average among courses. Unfortunately, our measures are usually confined to the four years during which the student is enrolled and qualities which can be measured by paper and pencil. In part, our deficiencies are methodological. We are unsure of appropriate research design, we are perplexed with problems of data collection, and we have difficulty in defining appropriate criteria.

Accordingly, one area of major contribution which should be made by offices of institutional research is to explore these methodological problems and help define appropriate solutions. And one of the most fruitful ways to refine otherwise crude methodological tools is to move directly into the conduct of the research.

It is helpful to occasionally look backward and with the benefit of such glances to see more clearly into the future. Looking backward, there is substantial evidence to suggest that offices of institutional research and institutional researchers have been preoccupied with the methodology and the conduct of input analyses. Indeed, the major theme of our 1966 annual forum was *input*. Certainly, it is in this area that institutional research has made its greatest contribution, and it is in this area in which college and university presidents, faculty members, and boards of trustees expect performance. The security to be derived by a *fixation* on input analyses to the exclusion of analyses of process, context, and output will likely be achieved at a substantial sacrifice not only to institutional researchers, but also to the colleges and universities as well as the systems of higher education which they serve.

FOOTNOTES

1. John E. Stecklein, "The Birth of a Profession." *Research on Academic Input*, Proceedings of the 6th Annual Forum of the AIR, 1966.

2. Samuel Baskin, "A Conceptual Framework for Institutional Research," *A Conceptual Framework for Institutional Research*, Proceedings of the 4th Annual Forum, May 1964.

3. Lewis B. Mayhew, "Imperatives for Institutional Research," *Research on Academic Input*, Proceedings of the 6th Annual Forum of the AIR, 1966.

4. John Henry Newman, *The Idea of A University*, (Edited by Martin J. Svaglic). New York: Rinehart & Co., Inc., 1960 edition.

5. Nathan M. Pusey, *The Age of the Scholar*, Cambridge, Mass., Harvard University Press, 1963.

6. Stanley O. Ikenberry, Eddie C. Kennedy, Juanita V. Field, "Effects of Reading, Study Skills Improvement, and Reduced Credit Load on Achievement and Persistence of Failure Prone College Freshmen: A Pilot Study," submitted to Dept. of HEW, U. S. Office of Education, Washington, D. C., November 1966.

7. Ernest Hayes, "The Influence of Course Load on College GPA." *College and University*, Vol. XXXVII, No. 3, pp. 251-3. Spring 1962.

8. Earl A. Koile, Vincent A. Harren, and Carolyn Draeger, "Higher Education Programs," Review of Educational Research, Vol. XXXVI, No. 2, April 1966, p. 250.

9. Samuel Baskin, op. cit., p. 1.

10. W. J. McKeachie, "Procedures and Techniques of Teaching: A Survey of Experimental Studies," The American College, John Wiley & Sons, Inc., New York, p. 312, April 1962.

4/2/49

CRITERIA OF EFFECTIVE TEACHING BEHAVIOR
IN AN INSTITUTION OF HIGHER EDUCATION

Richard R. Perry
Director, Office of Institutional Research
The University of Toledo

The Problem

Identification of teacher effectiveness is so complex that apparently no one knows today what "the competent teacher" is. The anonymity of the "competent teacher" has been the spur for countless research studies.[1] Gage stated that literature on teacher competence is overwhelming; so much so that even bibliographies on the subject are unmanageable. Although numerous studies are reported in the literature, few if any facts are firmly established about teacher effectiveness. There is no approved method of measuring competence which has received wide acceptance.[2]

The elusiveness associated with identifying effective teaching behavior has undoubtedly prompted the large number of studies in the field. The importance of such identification justifies any such reasonable research. Inadequacies in present evaluation devices and harm that can be accomplished by using inappropriate criteria beg for enlightening research to identify characteristics of effective teaching behavior.

One of the most serious aspects of the problem of identifying effective teaching behavior is that without such explicit identification evaluations which take place are suspect. Significant faults which are assigned to present methods of evaluation focus chiefly on the following inadequacies:

1. Criteria included in evaluations have not been warranted by adequate research.

2. Persons who do evaluation are criticized for their lack of expertness in the very field in which they are operating.

3. Evaluation of teaching behavior has not proven to produce high reliability in longitudinal studies, when total effectiveness of teaching behavior is considered.

The lack of conclusiveness of previous investigations has not diminished the zeal with which the results of such investigations are put forward. Perhaps the most useful result of all such examinations and experiments is to more clearly identify the problems experienced in trying to arrive at clear definitions of effective teaching. A most important consideration in such research is to understand that substantive evaluation can take place only in terms of explicit objectives. Until objectives are defined and agreed upon, evaluations tend toward spuriousness. However, a corollary to the establishment of objectives is the identification of criteria of teaching behavior which hopefully will elicit, or at least assist in, the attainment of teaching objectives. Even when a careful definition of desirable outcomes (objectives in teaching) is attained, it does not solve the criterion problem. Since the major problem in research on teaching behavior is that of criteria,[3] it seems that research on the identification of criteria which can be warranted for the evaluation of effective teaching behavior might be helpful.

Such attempts in higher education are not new. They have been increasing in frequency in the last five years. Research on the identification of warranted criteria received a good deal of impetus from the work of Ryans whose argument for such research indicates that there are good teachers and good teaching, and that characteristic behaviors associated with this teaching should be identifiable.

Although they may be identified, it can be assumed that not every teacher can possess all the "good" behaviors or characteristics; thus the goal of such research needs to be the identification of those criteria of teaching behavior which are critical. The identification of such criteria has been left often to the expert opinion or to administrative standards. The use of such authority has resulted in criteria proving unfruitful and of temporary value. The argument has gained weight that the place to look for characteristics of teaching behavior which result in effective teaching is in the behavior of teachers. Such reasoning suggests searching out clusters of behaviors associated with effective teaching.

A word needs to be said about the meaning of effectiveness. A single piece of research cannot hope to explore all the dimensions implicit in a concept such as effective teaching behavior. The majority of research studies in this area have focused on the assumption that in searching for teaching effectiveness, the research seeks for properties of the teacher. This viewpoint assumes that effectiveness is an attribute of the teacher. A further assumption is that the effectiveness is not seriously deterred by other variables, which establishes an hypothesis about the adaptability of a teacher to teaching situations.[4]

To assume that the effective teacher is one who can accomplish educational objectives with students, aside from other variables, is to recognize that effectiveness as a term may have several meanings because it is identified with several different teaching situations. There is no harm in using the term effectiveness if one recognizes that it is related to a set of particular conditions.

It seems that development of formalized evaluation procedures in numerous institutions of higher education is based on the assumption that there is a set of criterion behaviors which are appropriate to particular institutions.

A Brief Appraisal of Evaluation of Teaching Behavior

Evaluation of teaching seems to enjoy great attention in the popular and professional press, but one needs to remember that systems of such evaluation have been operative in colleges at least since the early 1920's. Some procedures have resulted in evaluations being given to deans or department chairmen who, in turn, are privileged to confer with faculty about the evaluations. Apparently, other systems of evaluation make it possible for the results of such procedures to be made known to salary and promotion committees, and some merely have the results made known to the professor.

It seems that none of these systems of evaluation is without criticism, and a few of these criticisms are helpful in the identification of basic faults in the evaluations. Major criticisms which are a matter of record in the minutes of faculty meetings at a private college indicate that:

1. The present procedure cannot be intelligently considered as evaluation of effective teaching but would be better named "poll of student opinion."

2. The present system does little to help in determining which faculty will be kept, lost, or attracted to the campus.

3. Those involved in the evaluation are not by education, experience, or responsibility qualified to make the evaluation they are asked to make.

4. That the current evaluation is obligatory upon the faculty member is a violation of faculty rights.[5]

The above comments represent a core of a faculty's concern about evaluation procedures.

There are other thoughts which are based on inadequacies in systems of evaluation. These seem to center on the following:

1. An institution will decide to provide for evaluation of teaching and will choose evaluation items from rating instruments which are already in use at other institutions.

2. An institution, or indeed an entire state system of higher education, will decide to honor outstanding teachers with cash prizes but will leave the identification of these outstanding teachers to the judgments of persons in positions of administrative authority or to impressionistic evaluations of individual faculty. The comment of one professor who found himself involved in a system of higher education providing for such identification indicated that, "even if you wanted to try out for an award you wouldn't know how to change your teaching. This whole reward set-up is too much like a beauty contest."[6]

3. An institution will make it possible for the evaluation of teaching to go on in one college or in one department and not in all of the departments or colleges on a campus. Thus, some faculty feel imposed upon while others feel deprived of the opportunity for evaluation.

4. The most serious concerns about the evaluation of teaching focus on answering the question: "Evaluation for what purpose?" This question has not been satisfactorily answered on a majority of campuses.

5. An additional area of major concern is finding a satisfactory answer to the question: "What criteria can be justified in the evaluation of a faculty member's effectiveness as a teacher?"

There is little question but what evaluation of a faculty member's effectiveness as a teacher takes place. Students, faculty colleagues, and the administration, if one should happen to be known to the administration, all comment in one way or another about the qualities of teaching exhibited by the faculty member. Implicit in all such evaluations is the concept that some faculty must be exhibiting behaviors in their teaching which are considered to be characteristic of effective teaching. Finding out what those behaviors are and determining a relative importance for each of the identified behaviors could be a first step in construction of a model or set of behaviors associated with effective teaching in higher education at any institution of higher education.

The University of Toledo's study on criteria of effective teaching centered on identifying effective teaching behaviors and determining their relative importance. There are numerous studies which produce interesting statistical results concerning reliability, correlations, and the results of factor analysis. Difficulties in some of these arise because of methods used in selecting criteria for evaluation instruments. Procedures which have established evaluation instruments by choosing criteria already in use at other institutions, without testing the warrantability of these criteria for the institution where they are to be used, leave something to be desired. Statistical analysis can be accomplished with responses given to any criteria utilized in any rating instrument, but the question remains as to the warrantability of criteria which are put to use in such procedures.

The University of Toledo Study

Background

Interest in effective teaching is not new to the University of Toledo, but in the last two years it has received increasing attention from the university faculty and student body. The administration of the university in the spring of 1964 announced the establishment of four outstanding teaching awards in the amount of $1,000 each. These awards, financed by the Alumni Foundation, are given to four faculty members each year in recognition of outstanding accomplishments in teaching at The University of Toledo. The College of Education simultaneously introduced structured evaluation procedures for its own faculty. The College of Education provided that at the end of each term faculty members could voluntarily request students to respond to an evaluation instrument which focused on the qualities of teaching in those courses taught by the individual professor. The evaluation instrument not only operated for the individual instructor but for the course as well. The criteria in the instrument resulted from the studied deliberations of a faculty committee of the College of Education. Results of the College of Education evaluation procedure are made known to the individual faculty member and to the salary and promotion committee of the College of Education.

With these developments the Office of Institutional Research at the University evidenced an interest in conducting a research study within the university community to identify criterion behaviors which could be warranted for use in the evaluation of effective teaching behavior in higher education.

The study was proposed to the deans of the colleges and the Faculty Conference Committee, all of whom endorsed it. An advisory committee to the Office of Institutional Research was appointed. The advisory committee consisted of a representative of each college appointed by the dean of that college. The proposed research focused on the central problem of evaluating effective teaching in higher education. That problem without question is the identification of criteria warranted for use in such evaluations, for unless criteria used in such evaluations can be demonstrated as warranted for the purpose at hand, they would be irrelevant.

In structuring the study the Office of Institutional Research at the University of Toledo made the following assumptions:

1. Criteria for the evaluation of effective teaching are related directly to the academic community in which they are to be used, and the place to look for these criteria, which are most appropriate for one institution, is within the academic community represented by that institution.

2. Criteria for the evaluation of effective teaching in higher education should be established as the result of consultation with those most directly concerned with such teaching; namely, students, the faculty, and the alumni of all an institution's colleges.

3. Students, faculty, and alumni should have opportunity to express their thoughts freely as to what separate actions they believe contribute to effective teaching, without their responses being limited by procedures which force them to select behaviors from a suggested list of such criteria which do not originate within their own community.

The First Phase

The University of Toledo began in the spring of 1965 and proceeded during the academic year 1965-66 with the first phase of the study, the second phase completed in the academic year 1966-67. The first phase contacted a stratified sample of faculty, students, and alumni to obtain free response identifications of behavior which contributed in the judgment of the respondents to the effectiveness of teaching. In order that this could be done and the data handled effectively, response instruments were designed to the configuration of a data card. Response instrument No. 1, with a personal data card, was mailed to a random sample of the student body stratified by college and class rank, to every member of the faculty of the University of Toledo, and to a random sample of alumni stratified by colleges from which they had received their degrees. Each potential respondent of the sample was given a personal data card and fifteen response instruments No. 1.

Thirteen thousand six hundred and forty-three (13,643) individual responses were received, identifying "effective teaching behaviors." These responses were received from 812 students, 166 faculty, and 665 alumni. This response resulted in replies from 10 per cent of the student body, 30 per cent of the faculty, and 8 per cent of the alumni degree holders. The average number of behaviors identified by students was 8.7; by faculty, 8.2; and by alumni, 6.8.

These 13,643 identified behaviors were then "read" by a jury group to identify duplications in behaviors. The jury group was looking for criterion statements with the same meanings but different wording. Examples are the two following responses.

1. "Ability to keep presentation of subject matter at a level comprehended by the student."

2. "Ability to present subject matter at student level."

Though the wording is slightly different in each statement, each can be valued as meaning the same as the other. The result of this reading process was to categorize 13,643 individual behaviors into sixty criterion behaviors. The reading procedure had one jury person read the statements, placing them in categories of sameness, and then had these categories checked by second and third jury persons; thus, questions were raised as to the appropriateness of the classification of any one of the criterion statements.

An additional result of this reading process was to identify six major categories of effective teaching behaviors. These six contained individual behaviors which grouped themselves into major behavior categories representing concentrations of similar kinds of behavior to permit their identification as major separate areas of teaching behavior. The identification of the individual criterion behavior and the clustering of these into the six major criterion behavior areas ended the first phase of the study.

The Second Phase

With the criterion statements available, the task was to obtain judgments of how warranted these were for the evaluation of effective teaching behavior. This task was accomplished by designing a response instrument in which the criterion behaviors were listed. The order of their listing was provided by a random listing of numbers supplied by a random number program from the University computer. The instruments provided for a response to the importance of each criterion from critical importance through no importance. Each respondent was able to categorize himself by checking appropriate spaces.

A sample of students stratified by college and class rank and a similar sample of alumni by college in which they had earned degrees was presented with the instrument with all faculty. Usable responses were returned by 756 students, 850 alumni, and 187 faculty. Returns resulted in replies from 7.5 per cent of the students, 8.6 per cent of the degree-holding alumni, and 35 per cent of the faculty. These percentages of the academic community seemed adequate in view of present research practices.[7] Weights of five, four, three, two and one respectively were assigned to the response areas of critical, above average, average, below average, and no importance. These data were coded into punched cards and processed for statistical analysis to establish rank orders and rank order correlations for selected categories of responses.

The Third Phase

The University of Toledo identified four outstanding teachers in each of the years 1964, 1965, and 1966. Responses of this group were obtained and processed for the same statistical analysis as for other selected respondent categories.

The correlation of the ranking of the criteria by the outstanding teachers with those of all other groups in the study tested the order of importance established in the study against the judgments of a "jury of experts." Seemingly, this is further justification for the warrantability of the criteria in the order established for them by the responses of the total group.

A Possible Weighting Procedure

A criticism often leveled at evaluation procedures is that each criterion is assumed to be of the same value. The warranting of criteria in this study provides for a value factor to account for the demonstrated differences in importance of each criterion. This value factor for each criterion was established by assigning the weighted raw score totals of all groups for each criterion to that criterion. For ease in computation and handling, weighted scores have been identified as decimal value factors. Such value factors permit an evaluation instrument including all or selected criteria from the study to be constructed. An Effectiveness Evaluation Scale could use criteria from the research in the following fashion.

> Sample Item:
>
> Check the term which in your judgment best describes your professor's characteristic teaching behavior.
>
> This professor demonstrates comprehensive knowledge of his subject.
>
> ____ Always ____ Most of the time ____ Occasionally
>
> ____ Very Seldom ____ Never

A student marking the space "Always" would be giving the faculty member a "5" on that item which when multiplied by its value factor of .732 would give him a score of 3.66 on this one item.

The sum of the products of the criterion ratings and the criterion value factors would produce an effectiveness score.

Findings

1. All rank order correlations between selected groups of respondents are different from 0 at the .01 level of confidence for individual criteria.

2. Sixty criterion behaviors associated with effective teaching at The University of Toledo have been established as warranted for evaluation of such teaching.

3. The academic community of The University of Toledo is agreed on the importance of the sixty criteria in the rank order which is established in the study.

4. A table of weights of importance has been established to account for the importance of each criterion.

5. Rank order correlations are different from 0 at .05 level of confidence for the major behavior categories between seventy-two of the seventy-eight selected groups.

Observations

Research on the effectiveness of teaching indicates promise in clarifying issues related to the evaluation of teaching. Such research can also help prevent the perpetuation of error in these evaluations or at least provide an analysis of a major problem in any evaluation--the identification of criteria to be used. This study seems to have done this for the present at The University of Toledo. An additional useful result of this study is the providing of a value weight for each criterion which could be used in an evaluation instrument so that some accounting of the differences in importance of criteria used in such evaluations may be accomplished.

The study reported here is apparently unique in that it provides a sample of one institution's total academic community with an opportunity to participate in consideration of criteria which may be used in evaluation of effective teaching. Apparently it is the only study in which the judgments of a representative sample of a complex academic community on such criteria have been tested against a jury of outstanding teachers in an institution.

Of course, significant problems remain in the evaluation of effective teaching. They are:

1. The competence of persons doing the evaluation.

2. The test of reliability of the criteria and procedures which can only be accomplished through longitudinal studies.

It seems, however, that a sound beginning has been established with the identification of criteria in this study.

Mention was made earlier that six of the rank order correlations obtained for the major behavior categories between the seventy-eight selected groups were nonsignificant at the .05 level. To search deeper into why there was a lack of significance at the 5 per cent level the raw data were re-run in a computerized t test program to ascertain where the significant differences in judgments of the groups which produced the nonsignificant correlations lay. An identification of these results is indicated for each of the pairs which produced nonsignificant correlations at the 5 per cent level.

1. All vs Engineering Students. The All group placed a significantly higher mean value of importance on Personal Characteristics, Community Campus Involvement, and Scholarly Orientation than did the Engineering Student Group.

2. All Students vs Engineering Faculty. The Engineering Faculty placed a significantly higher mean value of importance on Scholarly Orientation, while the All Student Group placed a significantly higher mean value on Evaluation Procedures.

3. All Alumni vs Engineering Faculty. Engineering Faculty placed a significantly higher mean value of importance on Scholarly Orientation than did All Alumni.

4. All Faculty vs Engineering Students. All Faculty placed a significantly higher mean value of importance on Scholarly Orientation than the Engineering Students, while the Engineering Students placed a significantly higher mean value on Evaluation Procedures.

5. All vs UCATC Alumni. No significant difference in means was established by the use of the t test.

6. All Faculty vs Law Alumni. Again as in the case of the All Group vs UCATC Alumni, the t test disclosed no significant difference in mean values of importance established by the two groups.

A limitation of the study in oversimplifying the grouping of the sixty criteria into only six major behavior categories as well as a lack of interval scale measures may be the cause of nonsignificance.

Because the College of Engineering was represented in four of the six nonsignificant rank order correlations and because in all four instances either Engineering Students or Engineering Faculty were represented, it was decided to run a t test on the values of importance established by the Engineering Students and the Engineering Faculty. The result of that test was that Engineering Students placed a significantly higher mean value of importance on Evaluation Procedures, while Engineering Faculty placed a significantly higher mean value of importance on Scholarly Orientation.

One might consider that students in general are more concerned with Evaluation Procedures than are faculty, but that in the special instance of Engineering Students it seems they are more concerned with such procedures than the general student body. All Faculty as a group seem more concerned with Scholarly Orientation than are students but, in this special instance, Engineering Faculty seem to place a greater weight of importance on Scholarly Orientation than the All Faculty group.

Prospects for the Future

Longitudinal studies on the criteria which have been established in this study would certainly add to validity, and the reliability of their application, if any, in the evaluation of effective teaching. The use of any or all of the established criteria in the evaluation of teaching could provide a reasonably sound basis on which to base studies concerning the competence of individuals undertaking evaluation. Perhaps their most fruitful possibility is in providing a framework within which to build longitudinal studies of the evaluation of effective teaching.

The University of Toledo study has resulted in a revision of procedures followed in the choosing of the outstanding teachers. The committee for choosing the outstanding teachers is composed of those faculty who have been identified in the past as outstanding teachers. It is their task to oversee procedures to be followed in choosing the next outstanding teachers. The committee has decided to use the first ten items in importance for the evaluation of effective teaching as identified in this study, which have been placed in an evaluation instrument prepared by the committee for choosing outstanding teachers. The instrument is made available to students, faculty, and alumni who indicate an interest in participating in the selection of and the nominating of an outstanding-teacher candidate.

Additionally, the College of Education has appointed a special committee to review its procedures and its instrument used in the evaluation of teaching in the College of Education. Thus, this study has been of measurable help in the identification of effective teaching at The University of Toledo.

CRITERIA RANK ORDER
ESTABLISHED BY ALL
SHOWING RAW SCORES AND VALUE FACTORS

CRITERION	RAW SCORE	VALUE FACTOR
35. Being well prepared for class	7757	.7757
27. Establishing sincere interest in the subject being taught	7412	.7412
20. Demonstrating comprehensive knowledge of his subject	7322	.7322
48. Using teaching methods which enable students to achieve objectives of the course	7321	.7321
2. Constructing tests which search for understanding on the part of the students rather than rote memory ability	7311	.7311
55. Being fair and reasonable to students in evaluation procedures	7274	.7274
10. Communicating effectively at levels appropriate to the preparedness of students	7264	.7264
47. Encouraging intelligent independent thought by students	7256	.7256
44. Organizing the course in logical fashion	7189	.7189
6. Motivating students to do their best	7139	.7139
41. Treating students with respect	7122	.7122
5. Acknowledging all questions to the best of his ability	6997	.6997
43. Being able to show practical applications of subject matter	6970	.6970
49. Rewriting and updating tests	6839	.6839
39. Being readily available for consultation with students	6724	.6724
36. Setting high standards of achievement for students	6646	.6646
42. Raising the aspirational level of students	6633	.6633
29. Recognizing his responsibility for the academic success of stud.	6568	.6568
17. Establishing good rapport with students in the classroom	6567	.6567
33. Patiently assisting students with their problems	6563	.6563
9. Having practical experience in his field	6520	.6520
60. Encouraging moral responsibility in students by his example	6483	.6483
24. Accepting justified constructive criticism by qualified persons	6479	.6479
22. Encouraging student participation in class	6461	.6461
15. Making written comments on corrected returned assignments	6452	.6452
1. Evidencing better than average speech qualities	6353	.6353
11. Identifying his comments which are personal opinion	6272	.6272
4. Engaging in continued formal study in his field	6251	.6251
57. Using more than one type of evaluation device	6219	.6219
31. Demonstrating a stable level-headed personality	6206	.6206
28. Taking measures to prevent cheating by students	6181	.6181

CRITERIA RANK ORDER
ESTABLISHED BY ALL
SHOWING RAW SCORES AND VALUE FACTORS (continued) p. 2

CRITERION	RAW SCORE	VALUE FACTOR
3. Providing several test opportunities for students	6153	.6153
54. Indicating that the scope and demands of each assignment have been considered carefully	6098	.6098
32. Returning graded assignments promptly	6036	.6036
56. Relating course material to that of other courses	6029	.6029
40. Displaying broad intellectual interests	5960	.5960
59. Exhibiting a genuine sense of humor	5959	.5959
18. Making an effort to know students as individuals	5894	.5894
53. Seldom using sarcasm with students	5887	.5887
21. Exhibiting an intelligent personal philosophy of life	5868	.5868
16. Presenting organized supplementary course material to students	5822	.5822
46. Earning the respect of his colleagues	5802	.5802
12. Challenging students' convictions	5753	.5753
51. Explaining grading procedures	5744	.5744
7. Explaining grading standards	5723	.5723
14. Announcing tests and quizzes in advance	5713	.5713
13. Utilizing visual aids to assist in creating subject matter achievement with students	5559	.5559
50. Presenting an extensive lucid syllabus of the course to students	5501	.5501
19. Inspiring students to continue for graduate study	5464	.5464
25. Sharing departmental duties with his colleagues	5311	.5311
23. Beginning and ending classes on time	5303	.5303
58. Being neatly dressed	5298	.5298
38. Being knowledgeable about the community in which he lives	5261	.5261
26. Having irritating personal mannerisms	4960	.4960
37. Involving himself in appropriate university committees	4588	.4588
34. Holding membership in scholarly organizations	4546	.4546
52. Being consistently involved in research projects	4321	.4321
30. Devoting time to student activities on campus	4278	.4278
45. Making appearances which assist programs of community organizations	4275	.4275
8. Publishing material related to his subject field	4191	.4191

FOOTNOTES

1 Gage, N. L., Address appearing in Proceedings, "Research Resume 1960", 16 Burlingame, California, California Teachers Association.

2 Biddle, Bruce J., Contemporary Research on Teacher Effectiveness, Holt, Rhinehart, and Winston, Inc., 1965, p. 2.

3 McKeachie, W. J., in Gage, N. L., editor, Handbook of Research on Teaching, American Educational Research Association, p. 1124 Rand-McNally, 1963.

4 Fattu, N. A., "Research on Teacher Evaluation," The National Elementary Principal, Vol. 63, No. 2, November, 1963, p. 19.

5 Minutes, Faculty Meeting, Antioch College, April 25, 1964.

6 Old Oregon, January-February, 1966, p. 13.

7 Holland, John L. and Richards, James M., ACT Research Reports, Oct. 1965, No. 8, p. 3.

STUDENT-FACULTY RATIOS: WHAT DO THEY MEAN?

John G. Bolin, Assistant Professor
Institute of Higher Education
University of Georgia

At first glance, student-faculty ratios appear to be a very simple mathematical matter. It is derived merely by dividing the total number of students by the total number of faculty members. And it is by this simple process that many institutions determine faculty needs. Chief administrators often base their overall plans and budget requests on this ratio as a matter of expediency.

Yet, the student-faculty ratio is a difficult factor to pin-down in the descriptive analysis of any university, and probably the most frequently misunderstood. One university president feels that because of varied faculty responsibilities a "faculty-student" ratio is "impossible to accurately compute, and is meaningless once it has been done." Nevertheless many administrators and most accrediting agencies consider the student-faculty ratio a major factor in the measurement of instructional quality of an educational institution. Inherently, therefore, it is an essential element in both administrative and instructional planning.

The chief complaints against using student-faculty ratios as measures of effectiveness consider the lack of a uniform formula for deriving the ratios, and to the lack of research from which conclusions can be drawn to determine optimum class size for any level of education. Both of these objections appear to be animadversions to the ratio or its use.

Regarding the first complaint, student-faculty ratios may be calculated in a variety of ways. In some instances headcount is the basis for computation, while in others the ratio is derived on an equivalent-full-time basis; also, there seems to be disagreement about whether to include graduate teaching assistants, research assistants, and part-time students in the tabulations. Since student-faculty ratios may be derived for a variety of purposes, a variety of computational methods is necessary. For instruction, unless otherwise stipulated, the standard practice seems to be to compare the number of equivalent-full-time students to the number of equivalent-full-time professional instructional staff. Deviations from this practice can be misleading and virtually useless.

The latter complaint regarding class size or faculty load is somewhat more complex. It is generally agreed, however, that large classes do not necessarily solve the problem of expanding enrollments.[1] Moreover, many educators feel that learning may be accomplished better in smaller classes, where exchange of information and ideas is not precluded by the number of students enrolled in class. Although studies which would support either point of view are limited, Anderson[2] and O'Shaughnessy[3] have shed some light on this argument.

Thus, while student-faculty ratios may not provide the micrometer by which quality may be determined, they can provide a yardstick to assist administrators and educators in the evaluation and improvement of their instructional program. A knowledge of the present status of class size and student-faculty ratios should be helpful to those working to build better institutions of higher learning. Hence, the purpose of this study is to help keep college and university officials informed about what is taking place in higher education with regard to expanding faculties and growing enrollments. I do not wish to debate the merits or demerits of the uses of student-faculty ratios; rather this investigation will be confined to analyzing the ratios as they were reported.

Specifically, this study endeavors to analyze the following factors as they relate to the student-faculty ratios reported by the colleges and universities included in the sample:

1. Means, medians, ranges.

What are the means, medians, and ranges of the student-faculty ratios in the various types and sizes of higher education institutions throughout the United States?

2. Type of control.

Are there any significant differences in the student-faculty ratios at publicly controlled colleges and universities, and those reported for privately controlled institutions?

3. Size of institution.

Are there significant differences in student-faculty ratios which could be attributed to size?

4. Geographic location.

Are there significant differences in student-faculty ratios related to geographic location of the colleges and universities?

5. State universities.

Are there significant differences among state universities according to the geographic region in which they are located?

6. Level of instruction.

Do instructional levels (lower division, upper division, graduate/professional division) vary in proportions of faculty assigned for instruction?

7. Type or level of authority.

By whom or what authority are the student-faculty ratios determined for the various types of institutions?

Procedure

A questionnaire requesting information about present student-faculty ratios and the authority determining them was sent recently to fifty major state universities (one in each state) and one hundred other colleges and universities of different sizes situated throughout the nation. The latter group included both privately and publicly controlled institutions. Responses were received from forty-seven state universities and seventy-two other schools and colleges.

The questionnaire used for this study contained only two questions. While one may criticize it for lacking sufficient definition or instructions for completion, the purpose was clearly implied in the first question which requested information relating to instructional personnel. Moreover, except for public institutions with student enrollments of 7,500 or more (excluding state universities), the number of responses distributed among the various sub-classifications was believed sufficient to allow for institutional variations in computing student-faculty ratios to be cancelled out between categories. Because of the lack of data regarding the large public institutions, no statistical comparison was made between private and public control in that specific category alone.

Analysis of the Data

While the data included in this study were limited by the lack of precision with which the terms were defined, they were sufficient to allow for the computation of arithmetic means, medians, ranges, and in some parts of the study for non-parametric statistical treatment.

In the initial part of the analysis all ratios were reduced to the number of students to one faculty member, and these scores were tabulated under the several sub-group headings according to type of control, student enrollment, and geographic location. After all scores were listed in their appropriate groups, the means were computed and the medians were located. The ranges of the categories were then given in order to provide a more complete description of each category.

Ultimately, where possible, the data were treated statistically with the Mann-Whitney *U*-test for uncorrelated data and the Kruskal-Wallis H-test for *k* sample cases. These tests were used to measure the differences in the distributions of student-faculty ratios among the various sub-categories identified in the study. The .05 level of confidence was accepted for identifying significant differences in the comparisons.

Intergroup Comparisons of Student-Faculty Ratios

The first analysis pertained to the relationship of control and enrollment to student-faculty ratios. In this analysis, the major universities were excluded from the tabulations because of the possible effects of their unusually large enrollments. To determine the influence of type of control upon student-faculty ratios, the total group was divided into the classical dichotomy of public and private. Each of the two types was then subdivided arbitrarily into three groups according to size. The classifications for the subgroups were listed as small, including institutions with an enrollment of under 2,500; medium, including those institutions having enrollments of 2,500 to 7,500; and large, including institutions with more than 7,500 students. Because of the lack of data pertaining to large publicly controlled institutions (other than the major state universities), no comparison is made between private and public colleges and universities in this category.

For the small institutions the mean number of students per faculty member was 13.0 for the twenty private institutions as compared with 19.2 (or about one-third more) for the twelve public institutions. (See Table I) The medians for the two types were 13.0 and 20.0 respectively, while the ranges of the two groups were 5.6 to 20.4 and 13.0 to 26.0. When the ratios of the two groups were treated statistically, the private institutions were found to have a significantly lower student-faculty ratio than the publicly controlled institutions. The Mann-Whitney *U*-test for uncorrelated data was computed to be 24.5, which when applied to the table of critical values, was found to be significant at the .05 level of confidence.

In the analysis of the two types as they were tabulated in the medium size category, the mean for thirteen private institutions was 14.6 students per faculty member, compared to the much larger mean of 20.1 for the nine public colleges and universities in the sample. The medians for the two groups were 12.0 for the private group ranging from 6.0 to 38.4 and 20.0 for the public group which ranged from 13.9 to 26.8. As above, this comparison indicated that the private institutions had a significantly lower student-faculty ratio than those in the public group. The *U* computed for this comparison, 21.5, was found to be significant at the .05 level of confidence.

To cross-check the results of the two preceeding comparisons between the public and private institutions, the two groups were combined and compared solely on the basis of control. Without regard to size, the mean for the sample forty-one private institutions was 14.4, while that for the twenty-two institutions in the public control category was 19.3. In this comparison the median of the private schools was 13.0, ranging from 5.6 to 38.4, as compared to a median of 20.0 and a range of 13.0 to 26.8 for the public colleges and universities. Since the samples comprised of the combined groups contained more than twenty cases, a *Z* value was obtained. The *Z* computed for this comparison was 3.99 which was significant beyond the .01 level of confidence.

To determine the effects of size upon the student-faculty ratios, the same institutions were redistributed with regard only to size. No consideration was given to the type of control. The distribution of cases among the three classifications were: small, thirty-two institutions; medium, twenty-two institutions; large, nine institutions. For this comparison the Kruskal-Wallis analysis of variance for non-parametric data was used to identify significant differences. The respective means for the small, medium, and large categories were 15.3, 16.7, and 17.5; their medians, on the other hand, were 15.0, 14.5, and 17.0. The ranges of the three groups were 5.6 to 26.0 for the small institutions, 6.0 to 38.4 for the medium size schools, and 8.3 to 30.0 for the large colleges and universities. The *H* value computed in the statistical treatment for comparing the differences in three groups classified by size was 2.67, which was not significant at the .05 level of confidence.

It is evident from these data that, on the average, privately controlled institutions of higher education have significantly lower student-faculty ratios than do their publicly controlled counterparts. It is evident also that size was not a factor in this difference. In the three comparisons where the type of control was the primary consideration, the privately controlled colleges and universities, without exception, indicated significantly lower student-faculty ratios; when the ratios were analyzed with size only as the primary consideration, however, no difference was found among the three groups.

Regional Differences

The second part of the study concerned the distribution of student-faculty ratios among the various institutions in the sample as they were classified by regional location. To make these comparisons each institution was placed arbitrarily into one of five geographic subdivisions of the United States. These included the Northeast, Southeast, Midwest, Southwest, and Far West. This phase of the study consisted of three analyses: regional differences excluding major state universities; regional differences including major state universities; and regional differences among major state universities only.

The first analysis of regional differences included a total sample of sixty-three colleges and universities. The total was subdivided without regard to size or type of control into the five regional categories. The numbers of institutions in each group included twenty-one in the Northeast; fourteen in the Southeast; seven in the Southwest; seventeen in the Midwest; and four in the Far West. The mean number of students per faculty member for the five groups ranged from 13.7 in the Far West to 19.9 in the Southwest. (See Table 2) Ranking second, third, and fourth respectively in terms of numbers of students per faculty were the Northeast with a mean of 14.0, the Midwest with 16.6, and the Southeast with 17.4. The medians for the groups were slightly different in rank-order and only reversed the ranking of the Far West, which had a median of 14.4, and the Northeast, whose median was 13.0. The Northeast also had the longest span of difference ranging from 6.0 to 38.4, while the Midwest had the shortest span, with a range of 11.9 to 26.0.

When these data were treated by the Kruskal-Wallis analysis of variance to test the differences among the groups, the H factor of 12.95 was found to be significant at the .05 level of confidence. Although differences between groups were not tested for significance, it is evident that there is significant variation in student-faculty ratios among institutions of higher education located in the five regions defined in this study.

In the second analysis of regional differences major state universities were included in the tabulations of the five regions. The numbers of cases in the individual categories were increased to thirty in the Northeast, twenty-three in the Southeast, fourteen in the Southwest, twenty-seven in the Midwest, and thirteen in the Far West. The means for the various regions in this distribution ranged from 14.6 in the Northeast to 19.5 in the Southwest. Ranking second, third, and fourth were the Far West with a mean of 15.7, the Midwest with 16.7, and the Southeast with 17.4. On the basis of the medians, the groups remained in the same ranking as that based on the means. The extent of the widest range was 6.0 to 38.4 in the Northeast, while the narrowest range, found in the Southwest category, extended from 11.9 to 26.0.

The Kruskal-Wallis H computed for this distribution was 6.6. This value was found not to be significant at the .05 level of confidence. From the results of this analysis it appeared that the state universities helped to offset the variation in ratios which was found to exist among the regions in the first analysis.

To cross-check the finding in the second analysis of regional distributions, tabulations which included only the state universities were tested for the significance of their variation. The number of cases included in this distribution were nine in the Northeast, nine in the Southeast, eight in the Southwest, ten in the Midwest, and eight in the Far West. The mean student-faculty ratios for these groups were 15.8, 17.5, 19.1, 16.8, and 16.4 respectively. (See Table 3) For the total group the mean ratio was 17.1:1. The medians were very close to the means, but again, the rank by median reversed the positions of Northwest and Far West regions. The largest range among these groups was 10.5 to 21.0 in the Southeast, while the smallest was 16.0 to 21.0 in the Southwest.

The H factor computed in this analysis was 5.9, which was not significant at the .05 level of confidence. With no significant difference among the regions where only state universities are included, it was increasingly evident that the major variations in student-faculty ratios by region lie within the smaller institutions.

Student-Faculty Ratios for Three Levels of Instruction

The data regarding student-faculty ratios, calculated according to level of instruction, was not sufficient for analysis in depth. Of the 107 cases included in the sample less than one-third (33) were able to make the necessary distinctions in their faculty assignments with regard to level of instruction; hence, these data were not treated statistically for significance of difference. However, the responses from the institutions completing this part of the questionnaire indicated a very definite trend of decreasing student-faculty ratios as the level of instruction was increased. The number of students per faculty member, as a rule, decreased approximately 33 percent between the freshman-sophomore level of instruction and the junior-senior level, the sophomore level of instruction and the junior-senior level. The decrease in student load was even greater in the graduate-professional level; at this level the student-faculty ratio was approximately one-fourth to one-third of that indicated for the lower division and about half as large as the ration indicated for the upper division of undergraduate instruction. Although this result should not be too surprising, it indicates a considerable differential in faculty load for those who are given specific instructional assignments in freshman-sophomore courses.

Prescribing Authority for Student-Faculty Ratios

The responses in this area of the investigation were arranged in six categories: Board of Trustees, State Board of Control (including Board of Regents), State Legislature, Administration, Other Source, and Not Prescribed. The data were not treated statistically, but the 107 responses were classified into Private, Public, and State University groupings. The percentages of the responses of the three groups were then calculated for each of the sources of authority, and comparisons were made only within the individual groups.

In the group of private institutions, only 34 percent had student-faculty ratios which were prescribed. Of these, 27 percent were derived from the Boards of Trustees, with the remainder of the prescription from the Administration. For the vast majority of private colleges and universities, no ratio was prescribed.

Student-faculty ratios in the public institutions were prescribed by several sources of authority, although there seems to be no central tendency for the prescriptions in this grouping. Nevertheless, in nearly half of these institutions ratios were not prescribed.

Among the state university group there also seemed to be no primary source of authority which determined the specific student-faculty ratio. Of those whose student-faculty ratio was prescribed, the most frequently cited authority was a state board of control or board of regents; this category, however, accounted for only seven of the forty-four cases in the sample. Slightly over 68 percent of the universities had no prescribed ratio.

Perhaps the most significant finding in this division is that 63 percent of the colleges and universities in the sample received no predetermined student-faculty ratio from any specified source of authority. A number of institutions indicated that the size of the faculty was dependent upon legislative budget allocations, or as in the case of the privately controlled institutions, upon gifts and contributions to established funds. Hence, it seems that the financial condition of the institution may be the chief factor in determining the quality and number of faculty a college or university can contract to instruct its students.

Summary and Conclusions

The study was undertaken to provide current information regarding the student-faculty ratios presently maintained in the various types and sizes of colleges and universities across the nation. The project was designed to investigate differences in student-faculty ratios among colleges and universities classified by type of control, student enrollment, and geographic location. The investigation was extended to determine the sources of authority which prescribed the student-faculty ratios. This study was limited both in the size of the sample and in the precision with which student-faculty ratio was defined in the questionnaire. Nevertheless, the number of responses and the particular efforts made by the respondents to insure the clarity of their information provided sufficient data to make the analysis possible. It should be kept in mind, however, that it is within these limitations that the conclusions of the study have been drawn.

The findings were as follows:

1. Student-faculty ratios in privately controlled colleges and universities are, on the average, smaller than in comparable public institutions.

2. Excluding the major state universities, institutional size, according to student enrollment, does not significantly influence the magnitude of the student-faculty ratio; that is, neither the larger nor the smaller institutions can be expected to have a smaller ratio than the other.

3. The variation in student-faculty ratios among the geographic regions is significant when major state universities are not included. The ratios were found to be smaller in the Far West and Northeast, and higher in the Southwest and Southeast. When the state universities were included in the tabulations, no significant difference existed among the various regions, although generally the lower ratios were still found in the Far West and Northeast, with the higher ratios in the Southwest and Southeast.

4. There were no significant differences among state universities distributed regionally.

5. Student-faculty ratios are virtually impossible to compute by level of instruction because of the overlapping instructional assignments of faculty members. In spite of the limited response to this part of the questionnaire, there was an indication of a substantial inverse relationship between level of instruction and student-faculty ratio.

6. No single agency or individual appears to be a major source for prescribing student-faculty ratios. In fact, in almost two-thirds of the cases in the sample, the student-faculty ratio was ostensibly a product of the institutional budget, while in slightly over one-third of the cases a variety of agencies, boards, or individuals predetermined what the ratio would be.

From these findings it seems evident that the instructional load of faculty members in private institutions, with regard to numbers of students, is generally less than that of their fellow educators in public colleges. No difference in the average student load, however, can be expected either in the smaller or larger institutions when they are classified only by size of student enrollment. And in view of the generally larger student loads of faculty members in colleges and universities in the South, a greater effort must be made to increase the faculties if the Southern institutions are to reduce their student-faculty ratios to the national averages. Finally, while there is an ostensible lack of prescribed or predetermined student-faculty ratios among colleges and universities, perhaps some general guidelines could be established by the administrator to provide for better planning and faculty assignment. These guidelines may help to more equitably distribute faculty loads within an institution.

This has been only an exploratory study of student-faculty ratios. It would seem advisable, therefore, that further study be undertaken with regard to the ratios presently existing in the various areas of instruction, or perhaps even by courses. Ultimately, the study of student-faculty ratios may provide a more comprehensive understanding of faculty responsibilities and workloads as they are distributed among and within modern institutions of higher education.

TABLE 1
STUDENT-FACULTY RATIOS BY TYPE OF CONTROL*

Private VS. Public Colleges And Universities
(Excluding Major State Universities)

		Private			
$\bar{X}$ =	13.0	14.6	14.4		
MDN =	13.0	12.0	13.0		
Range =	5.6-20.4	6.0-38.4	5.6-38.4	$\bar{X}$ =	16.1
		Public		MDN =	16.0
$\bar{X}$ =	19.5	20.0	19.3	Range =	5.6-38.4
MDN =	20.0	19.0	20.0		
Range =	13.0-26.0	13.9-26.8	13.0-26.8		

STUDENT-FACULTY RATIOS BY SIZE ONLY

$\bar{X}$ =	15.3	16.7	17.5	16.1	17.1
MDN =	15.0	14.5	17.0	16.0	17.1
Range =	5.6-26.0	6.0-38.4	8.3-30.0	5.6-38.4	10.5-22.0

*All figures indicating means, medians, and ranges represent numbers of students to one faculty member.

TABLE 2
STUDENT-FACULTY RATIOS BY REGION

(Excluding State Universities)

$\bar{X}$ =	14.0	17.4	19.9	16.6	13.7	16.1
MDN =	13.0	17.5	20.0	15.0	14.4	15.5
Range =	6.0-38.4	7.0-30.0	7.7-26.8	11.9-26.0	5.6-20.5	5.6-38.4

STUDENT-FACULTY RATIOS BY REGION

(Including State Universities)

$\bar{X}$ =	14.6	17.4	19.5	16.7	15.7	16.5
MDN =	14.0	18.0	19.8	16.0	15.0	16.0
Range =	6.0-38.4	7.0-30.0	7.7-26.8	11.9-26.0	5.6-22.0	5.6-38.4

TABLE 3
STUDENT-FACULTY RATIOS BY REGIONS

(Major State Universities Only)

$\bar{X}$ =	17.5	15.8	19.1	16.8	16.4	17.1
MDN =	18.0	16.0	19.1	16.5	15.5	15.5
Range =	10.5-21.0	12.6-21.0	16.0-21.0	13.0-20.0	13.0-22.0	10.5-22.0

FOOTNOTES

1. Joseph R. Cammarosano and Frank A. Santopolo, "Teaching Efficiency and Class Size," School and Society, September 27, 1958, pp. 338-341.

2. Kenneth E. Anderson, "Relationship Between Teacher Load and Student Achievement," School Science and Mathematics, June, 1950, pp. 468-470.

3. Louis O'Shaughnessy, "The Size of College Classes and the Percentage of Failures," School and Society, September 14, 1929, pp. 374.

PROFESSORIAL ATTITUDES TOWARD THEIR WORK ENVIRONMENT AT THE UNIVERSITY OF MICHIGAN

Donald R. Theophilus, Jr.
Special Assistant to the Vice President
University of Washington

The United States is committed to, and is betting on, higher education to help shape its future. Yet, as a national news magazine (Time, 1966) stated recently in a cover story, the outcome of the bet "depends on solitary teachers in secluded classrooms...and the number of bored, hostile, and inadequate college teachers adds up to something between a serious concern and an outright scandal." It is not possible to assess the total cost of this dissatisfaction, but at the very least, some substantial fraction of the potential or talent represented by faculty members is not being translated into action.

The purpose of this study was to measure professorial attitudes toward selected aspects of their work environment, for only by understanding what professors consider to be in their best interest can academic administrators begin to unlock the reservoir of energy and creativity within faculty resources.

The study utilized a twenty-page questionnaire sent to the three professorial ranks and academic administrators at The University of Michigan in the spring of 1966. More than 1,500 responses and a 73 per cent return were received. The study was sponsored and financed by the Office of the Vice President for Academic Affairs at The University of Michigan. A complete report of the project is available in Theophilus, (1967).

Related Research

There is a growing volume of research by social scientists on the topic of work motivation. Much of this is within the area of business and is classified under organizational behavior. Major theoretical and empirical efforts in this area would include the Harvard Business School studies of Elton Mayo, Fritz Roethlisberger, Abraham Zaleznik, Paul Lawrence, and others; and the Michigan studies of Rensis Likert, Robert Kahn, and Daniel Katz. Other contributors of note include Frederick Herzberg and his colleagues, Chris Argyris, W. F. Whyte, and Douglas McGregor.

A substantial amount of this research relates to an investigation of "job attitudes," and the measurement of these attitudes has been extensively developed and applied. It is disturbing, however, that the empirical findings about work motivation have been far from conclusive.

Although there is much to criticize about current research of work motivation in business organizations generally, it is not as deplorable as within academe where studies of work motivation are virtually nonexistent. It is astounding that the methods of social research have been applied by university professors to almost every important American institution except their own (Caplow and McGee, 1958).

Essentially there are only three major books that have touched upon various and certainly limited aspects of faculty motivation. These are Wilson (1942), Lazarsfeld and Thielens (1958), and Caplow and McGee (1958). A variety of studies over the last ten years, including many by directors of institutional research at universities, has provided much of the empirical material now available. These major research projects include Russell (1962), Gustad (1960), and Stecklein and Lathrop (1960).

Other studies have been conducted within individual institutions or among groups of institutions. Of these studies, many have been concerned solely with faculty personnel policies or some other narrower aspect of the problem. Much of this institutional research has not been published. In summation, the empirical and theoretical study of work motivation and work attitudes in higher education is barely in its infancy.

Study Objectives

Today's multiversity is not the single community of scholars as in times past. Rather it is characterized (Pace and Baird, undated) as "an assortment of parts and pieces loosely held together by people and programs having roughly similar but by no means common goals." Each of these "parts and pieces" may well exert a differential impact upon the work attitudes of its professors, depending upon such factors as the environment of the organizational unit and the professors in it. On the other hand, despite the existence of differences in atmosphere and personnel between, say, engineering and education, music and medicine, professors may nevertheless share the same kinds of attitudes about their work environment. The primary objective of the study was to check these alternative speculations and to test the hypothesis that professorial attitudes about work environment differ significantly among eighteen major organizational units at The University of Michigan.

A basic concept dealt with deprivation, which was defined as the discrepancy between the importance a professor attached to a potential incentive within the work environment and his satisfaction with it. In the course of testing the prime hypothesis, it was also hoped to: (1) gain some understanding about which environmental incentives were most important and potentially satisfying, and caused less deprivation at a given point in time within each of the units tested; and (2) determine the effect of selected background characteristics (such as rank, age, salary, level of education, etc.) on professorial perceptions of the importance, satisfaction, and deprivation relating to these potential incentives.

A secondary objective was to test the hypothesis that across the several organizational units of the university there are: (1) significant differences between professors' and academic administrators' ratings regarding the relevant weight to be applied to selected criteria for promotion and salary increases; and (2) significant differences between professors' and academic administrators' perceptions of the importance and satisfaction attached by professors to the potential incentives.

Basic to this search for data was the belief that good communication is at the heart of any effective incentive system, and to prescribe for individuals it is necessary to diagnose and understand the desires and needs of the individuals concerned.

Methodology

The study utilized measurement procedures similar in form to those included in studies and questionnaires used many times with profitable results by researchers at The University of Michigan's Institute for Social Research and elsewhere. That a large "n" was used added confidence in the reliability and validity of the data. Filter means programs and t-tests were used extensively, and the pattern of statistical analysis was felt to be quite refined.

Two major classifications of incentives were chosen for response and analysis. Material, tangible, personal incentives were those that specifically could be offered to an individual; for example, physical facilities, salary, research funds, personnel services, and fringe benefits. Non-material, intangible, non-personal incentives were those factors that usually could not be offered specifically to an individual and that often could be known or recognized by their absence. Examples are administrative concern for personal welfare and career development, communication and participation in making decisions and formulating policy, independence and freedom, and leadership. In all, ten major incentives or potential satisfiers were tested.

Details concerning the instrument, population, mailing and returns, coding, tabulation, and computer analysis are found in Chapter Two of Theophilus (1967).

Major Findings

Impact of Organization

The data supported the hypothesis that professorial attitudes of satisfaction were significantly different among the units. Organizational units did exert a differential impact on professors' attitudes toward their work environment. Unfortunately there are virtually no known comparable data that test the phenomenon of attitude difference among academic units. Several broad areas of existing research seem relevant, however, and the inferences are that the differences in attitudes among units are: (1) attributable to differences in the environment of the unit itself (administrative policies and practices, traditions, peer group influence, etc.); and/or (2) attributable to personality types that gravitate toward a certain discipline or field. Because of cost the extensive testing on satisfaction was not conducted to determine if the importance and deprivation felt by professors also differed significantly among units. Data from five of the statistical tests indicated, nevertheless, that approximately the same number of significant differences among the units were appearing for perceptions of importance as for satisfaction.

Impact of Background Characteristics

The twelve tested characteristics had comparatively little effect on professorial attitudes of satisfaction, thereby supporting the conclusion that a professor's attitudes of satisfaction in larger measure depended upon the unit to which he belonged. Those characteristics having the most effect were six in number, respectively: salary level, number of years employed by The University of Michigan, level of education, rank, age, and duties or responsibilities. Data on the six most influential characteristics were compared with job-attitude findings about industrial workers. Professorial attitudes of satisfaction did not seem to differ markedly from those of industrial workers when related to age, length of service, and salary.

How Professors Felt About Their Work Environment

There was a wide difference in professorial feelings about the ten classifications of incentives tested. The findings indicated that certain potential satisfiers were much more influential than others in affecting the attitudes of professors. Deprivation (the discrepancy score between importance and satisfaction) was considered the most meaningful measure, and the greatest sources of deprivation were respectively associated with the categories of "Funds for Professional Development," "Leadership," "Income," "Communication and Participation," "Personnel Services," "Administrative Concern for Personal Welfare," and "Physical Facilities." See Table 1. While professors did not think material incentives were nearly so important as the non-material, they nevertheless expressed greater dissatisfaction about them.

As a group, the three professorial ranks at The University of Michigan seemed to be comparatively well satisfied with their working environment in the spring of 1966. The overall mean of satisfaction for Section I of the questionnaire was 3.528 on a five point scale, which placed the score just more than half-way between the categories "Neither Satisfied nor Dissatisfied," and "Satisfied." The average discrepancy score between satisfaction and importance for professors was just above half a point at .589. Neither figure indicated university-wide dissatisfaction or low morale on the part of professors.

TABLE 1

PROFESSIONAL ATTITUDES ABOUT THEIR CONDITIONS OF ACADEMIC SERVICE

	Satisfaction			Importance			Discrepancy Score	
Sub-Sections of Section I in the Questionnaire	Mean	Rank From Most To Least Sat.	Est. Std. Dev.	Mean	Rank From Most To Least Imp.	Est. Std. Dev.	Between Means of Sat. & Imp.	Rank From Lowest To Highest D.S.
Fringe Benefits	3.838	3	.630	4.081	8	.603	.243	1
Indep. & Freedom	4.388	1	.627	4.652	1	.421	.264	2
Academic Colleagues	3.922	2	.707	4.326	2	.495	.404	3
Physical Facilities	3.164	10	.617	3.723	10	.502	.559	4
Admin. Concern	3.656	4	.869	4.235	3	.464	.579	5
Personnel Services	3.457	6	.768	4.086	7	.530	.629	6
Commun. & Particip.	3.349	8	.648	3.981	9	.487	.632	7
Income	3.475	5	.808	4.215	4	.561	.740	8
Leadership	3.413	7	.952	4.184	5	.529	.771	9
Funds for Prof. Dev.	3.251	9	.734	4.107	6	.491	.856	10
Section I - Question	3.528	..	.526	4.117	..	.343	.589	..

Scale for Response

A How satisfied are you at present with each aspect of your work situation at The University of Michigan?					B How important would each aspect be to you if you were evaluating an academic position elsewhere?				
Very Dissatisfied (1)	Dissatisfied (2)	Neither Satisfied Nor Dissatisfied (3)	Satisfied (4)	Very Satisfied (5)	Very Unimportant (1)	Unimportant (2)	Neither Important Nor Unimportant (3)	Important (4)	Very Important (5)

Among units there was a marked difference in the mean of attitudes, ranging from 3.847 to 3.083 for satisfaction, and from .172 to 1.093 for deprivation. There was very little difference in the overall feelings of importance among the units. Analysis yielded three important measures about the units: satisfaction, deprivation, and the degree to which improvement was desired. A rank difference correlation for the three combinations of rankings was above .90 for the units.

Comparison of Selected Administrative-Professorial Perceptions

Significant differences were found between professors' and academic administrators' ratings regarding the relevant weight to be applied to selected criteria for promotion and salary increases. Five comparisons were made and labeled: downward communication, upward communication, professorial intra-personal conflict, administrative intra-personal conflict, and administrative-professorial inter-personal conflict. The measure of faculty intra-personal conflict was perhaps the most interesting and evaluated how professors thought administrators weighed the criteria for promotion and salary increases compared to how they thought administrators ought to weigh them. All but one of the tests showed a significant difference, indicating that professors were extremely unhappy about how they thought they were being evaluated. Furthermore, administrative-professorial perceptions of what incentives professors thought satisfying and most important were found to be significantly different in nearly 40 per cent of the tests.

Implications for Administrative Action

Since the study was limited to one university at a single point in time, only additional research will determine if the findings are generally applicable. The findings from the more than 2,400 tables produced do have extensive implications for administrators at The University of Michigan, however, and should raise serious questions in the minds of academic administrators elsewhere.

Two examples, one concerning organizational units and the other individuals, illustrate how administrators might make use of the findings. First, if professorial attitudes are significantly different among the units as the data indicate, then the most productive point of departure for making the university a more stimulating place in which to teach and research may be to deal with organizational units on an individual basis, rather than with the total university environment *en masse*. Perhaps more attention should be given to a flexible system of organizational rewards and leadership for each unit as the need dictates. Harold Leavitt has argued, for example, that large organizations cannot be managed as "unified wholes" with a single managerial strategy being equally appropriate for all of its varied sub-sections. Leavitt also notes:

> If this interpretation is correct, we should be seeing less and less uniformity in managerial practice; more of a class system, if you like--though a fluid one--in which the rules governing everything from the hours of work to methods of evaluation, to systems of compensation may vary from one group to another within the same parent organization. And, further, the many variations will be understandable in large part if one looks at the task that each group is trying to accomplish, and secondly at the tools, psychological and technical, that are at the moment available for working on those tasks.

Moreover, of further relevancy are some of the theories of Chris Argyris, which have been described (Gellerman, 1963, p. 81) in this manner:

> The ideal organization ...is not only one which is flexible enough to shift power downward or upward, as the occasion requires, but

> one in which the entire organization assesses its problems and selects the power distribution that seems most appropriate to the times. Argyris is proposing, in other words, to serve both the needs of the organization (by versatility in operating styles) and the needs of the individual (by giving him a voice in selecting whichever style is to be used).

The second example, which regards individuals, concerns data that reveals the early thirties and forties as the periods of greatest deprivation for professors. Department heads could be alert to the possibility that these periods in a professor's life might be ages of unusual stress. Anticipation of these periods could allow for preventive measures before morale and perhaps effectiveness are damaged. Face-to-face interaction between the professor and other key people in his department and college would be even more important at this time. At least several times annually, the professor and those who are most influential in his work and career should take the opportunity to review his work and its future direction. Sabbatical leaves usually occur at regular intervals, yet in terms of career cycles and stress, it might be best to have more frequent sabbaticals in the middle years and fewer later on.

The Challenge to Academic Administration

The professor who works in a multiversity finds himself in an environment consisting of a blend of people, working conditions, and the work itself. This complex of factors impinges on the professor during every working day, shaping his attitudes and forming his behavior. He may not be a prisoner of this motivational environment, but the findings of many studies indicate that too often this environment blocks his inclinations and drives him into dissatisfaction and negativism, at a time when the need has never been greater for effective teachers and individual fulfillment for both student and professor.

One of the most important tasks of academic administrators is to stop this loss of potential and talent and to help turn the working environment into a stimulus which will release energy and creativity. To achieve this goal will require soul searching on the academic administrator's part, since the findings of this and other studies would indicate that he has created many of the unfortunate aspects of the working environment himself, although often inadvertently (Gellerman, 1963, p. 93).

Many academic administrators are insensitive to psychological advantage--the individual's private notions about what constitute his own best interests. They cannot assume that their view of the work environment is shared by professors. Administrators with their "birds-eye" view of the organization can easily misjudge the perspective from which the individual professor or professors in a school or college judge an action. The professor behaves as if his environment were real, and the administrator acts upon the basis that his perceptions of the environment are real. Much administrative-faculty conflict results from this failure to realize that the other person sees his world in a different light. The challenge for the academic administrator is to accept the difficult responsibility of understanding rather than dismissing the attitudes of the professors with whom he must work. It lies in recognizing that he is more capable of changing his own habits than of changing those of fellow workers.

Academic administrative policy at all levels of the organization must make sense in light of the academic staff member's perception of his environment and what he regards as his own advantage. As Gellerman notes (1963, p. 200), "When that perception is understood, then perhaps misunderstanding can be cleared up, diverging points of view can be accommodated, or at least the lines of disagreement can be clearly stated." Effective communication requires knowledge of what professors want to hear, since this greatly affects what they will hear. Therefore, the attitudes of the individuals who make up the organization, as well as the environmental forces that shape those attitudes, must be intimately understood if academic administrators are effectively to guide the university toward its goals. In this context, it is remarkable that so few schools have systematically attempted to obtain from faculty members their opinions about work and working conditions.

The Need for Periodic Study

The foregoing observations strongly suggest that the collection of professorial attitudes should be a periodic, carefully chosen, conscious procedure which would enable administrators in higher education to more intelligently guide and make decisions about priorities. The use of assumption and guess needs to be superseded. For example, Likert (1961) feels it is difficult, if not impossible, to build and maintain effective organization without routinely gathering and analyzing causal, intervening, and end-result variables. As examples of causal variables for industry, he lists organizational structures and objectives, management practices and behavior, capital investments, and needs and desires of members of the organization. Intervening variables include personality, attitudes, and communication. End-result variables run the gamut from production to grievances, absence and turnovers. Furthermore, Likert (1961) maintains that instruments to measure these variables are available or could be developed by existing methodology. He also believes that analyses of these variables could be made which would show why the existing conditions occurred, what changes of different conditions are likely to yield better results, and how best and most effectively to go about producing changes in the desired direction. These are the kinds of questions that institutional researchers must begin answering for the future.

In summary, effective university planning requires an intimate and first-hand knowledge of how professors perceive their job environment (and their work). Academic administrators in conjunction with institutional researchers and social scientists must make this a continuing field for study. Whether such research leads to greater satisfaction (and less deprivation in work), to greater productivity, and to a greater release of talent and potential will depend, however, as much on the imagination, ingenuity, and skill of academic administrators at all levels, as on the social scientists and institutional researchers.

REFERENCES

Caplow, T., and McGee, R. J. The Academic Marketplace. New York, 1958.

Gellerman, S. W. Motivation and Productivity. New York, 1963.

Gustad, J. W. The Career Decisions of College Teachers. Research Monograph Series No. 2. Atlanta, 1960.

Lazarsfeld, P. F., and Thielens, W. Jr. The Academic Mind. Glencoe, Illinois, 1958.

Leavitt, H. J. "On Human Organizations," Harvard Business Review, XL, No. 4, 1962, 90-98.

Likert, R. New Patterns of Management. New York, 1961.

Pace, C. R., and Baird, L. "Patterns of Attainment Related to the Environmental Press of College Sub-Cultures." Unpublished monograph. University of California at Los Angeles, (Undated).

Russell, J. D. "Faculty Satisfactions and Dissatisfactions," The Journal of Experimental Education, XXXI, No. 2, 1962, 135-139.

Stecklein, J. E., and Lathrop, R. L. Faculty Attraction and Retention. Minneapolis, 1960.

"Teaching: To Profess with a Passion," Time, (May 6, 1966), 80.

Theophilus, D. R. Jr. "Professorial Attitudes Toward Their Work Environment at The University of Michigan: A Study of Selected Incentives." Unpublished doctoral dissertation, The University of Michigan, 1967.

Wilson, L. The Academic Man. New York, 1942.

THE ACADEMIC STAFF QUESTIONNAIRE AS A RESEARCH TOOL

Paul F. Mertins
Research Associate, Office of Institutional Research
University of Michigan

The purpose of this presentation is to explore the academic staff questionnaire from the standpoint of a research tool. We at the University of Michigan have learned a great deal from a survey of academic staff attitudes which may be of great importance to other institutions. Since conducting this study, we have had several inquiries from other institutions planning similar studies. Hence, we concluded that many institutions could benefit from our experience. In writing this paper it became obvious that observations arising from the results of this study were not unique to the analysis of this particular area; rather, they could be applied to most research projects using a questionnaire as a gathering device.

Administrators are pressured more and more to make decisions based upon the very latest information. They are less willing to make decisions on outdated historical data and turn increasingly toward the current survey research approach in gathering information.

The following discussion on planning, problems, recommendations for improvement, and costs should be of specific interest to anyone considering a survey of faculty attitudes. It will be most helpful, in a general sense, as a guide to people who wish to undertake an original survey with the use of a questionnaire as a data-gathering device. In undertaking this study, we were fortunate to have available to us the experience and expertise of the Institute for Social Research of the University of Michigan.

The essential first step in undertaking any research project is to develop objectives and/or hypotheses. This not only prevents researching for the sake of research; it also defines more clearly the problem of analysis and report format. For us, the practical aspects of this lay in deciding upon a target population and what variables we would have to identify to analyze the data in proving or disproving our hypotheses. In the case of the academic staff questionnaire, the two hypotheses were as follows: (1) The importance, satisfaction, and deprivation (discrepancy score between importance and satisfaction) attached by professors to certain potential incentives will differ significantly among the many organizational units of a university, and (2) Across the many organizational units of a university, there will be the following: (a) significant differences between professors' and academic administrators' ratings regarding the relevant weight to be applied to selected criteria for promotion and salary increases, and (b) significant differences between professors' and academic administrators' perceptions of the importance and satisfaction attached by professors to certain potential incentives.

Thus, if we were to explore faculty rank by school and college, we would have to send questionnaires to all faculty members rather than use a sampling technique. It was determined that we did not have enough total population in some areas to make the sample meaningful, since some of our smaller colleges have very few people of professorial rank. After drawing a sample of this very small total population and then anticipating a very low return, it was determined that in some areas we might receive only one or two individual responses. This would be too small a response for statistical analysis. Another factor we could not anticipate was the percentage of return. The only related information we had was on a similar study conducted at New York University in which the return was 45 per cent. We at the University of Michigan received a total return of 60 per cent with a response of 73 per cent in the professorial areas.

We decided to send questionnaires to part-time faculty members such as Teaching

Fellows and Research Assistants. Since we were seeking responses from both full-time and part-time teaching personnel, we felt it would be logical to extend the study to include all staff members holding academic appointments. If a decision is made to restrict the sampling, it must be ascertained that it is possible to sort out the desired population mechanically by rank or other identifying characteristics. It is almost impossible for large institutions to perform this task manually.

It is highly desirable to employ a critical path scheduling technique once the type of gathering device, the sample (or n) to be surveyed, and the basic format of the type of analysis have been determined. The advantages of this type of planning are numerous. For example, it will immediately highlight the timing of the distribution and retrieval of the data as critical factors. A scheduling device such as this allows for seeing at a glance which functions are parallel and which overlap. It prevents misdirecting available manpower into areas in advance of their desired completion at the expense of functions critical to the timing of other functions. As functions were completed on the schedule, they were color-coded, and a line was drawn perpendicular to the time axis and moved along in conjunction with the current date. At a glance, this process immediately identifies functions which are ahead of schedule and functions which are trailing behind.

If such a study were to be conducted again, one of the most important considerations would be to estimate that the time allotted for computer processing and analysis would be much greater than had been anticipated. It would also be advisable to direct some effort into these functions as soon as possible in order to plan and revamp necessary programs. It is not wise to begin programming or reprogramming existing programs until a final decision has been made on the form in which the input data are to be used. Generally, when a gathering device such as the questionnaire goes to the printer, the input data are finalized. There is a tendency not to realize that the analysis could start at this early point in the processing; therefore, our critical path schedule could be modified to indicate this.

The budgetary and administrative red tape of any data processing service group at a large institution is always time-consuming. These functions have no real relationship to the data's availability for processing. Programming and the modification of canned programs should start as soon as possible, however, because of the necessity of testing program modifications and the availability of computer "turn-around" time.

If financial resources are limited in conducting a study of this nature, developing frequency responses should be a primary goal. This type of analysis yields much more information, and arithmetic means can be developed by hand from the frequency responses for selected groups. Programs that can develop both simultaneously by machine constitute the most desirable method.

In our particular case, the questionnaire was developed basically for the teaching staff. When we decided to send it to all academic staff members, the questions were modified to include respondents other than teaching faculty members. This modification had a tendency to weaken the questionnaire. If possible the questionnaire should be developed for the specific groups involved, either by sending special questionnaires to selected groups or by sectioning the questionnaire so that the responses to particular parts will emerge more meaningfully from the group primarily involved in these activities.

The attached cost exhibit indicates which costs would vary with the n. This information will allow you to estimate the approximate expense for a similar study, if you have an estimate of the population to be surveyed. Office of Institutional Research salaries represents our closest estimate of the wages paid to O.I.R. staff members involved in the study. Another way of estimating the cost of this item is to allow for the salaries of two full-time employees for a twelve-month period. These costs do not include time spent by administrative staff in deciding what information is to be gathered and how it is to be disseminated; nor do they include the time required for writing a final report.

ACTUAL AND ESTIMATED COST

ACADEMIC STAFF QUESTIONNAIRE

Cost	Item	Quantity
$ 109.80	9 ½ x 12 ½ Envelopes	6,000
96.00	9 x 12 Envelopes	6,000
45.00	#10 Reminder Envelopes	6,000
28.00	Covering Letter	6,000
28.00	Reminder Letter	6,000
13.82	Fold and Stuff Reminder Letter	6,000
20.91	Label Reminder Letter	6,000
95.13	Address and Label Sending and Return Reply Envelopes	12,000
558.67	Questionnaires (Typesetting)	19 Pages
746.91	Questionnaires (Printing)	6,000
70.89	Labels (DPC)	20,900
1,108.65	Postage (sending 9 ½ x 12 ½ Envelope with Questionnaire)	5,543
520.43	Postage (V.S.P.O. First Class Permit)	2,600
277.00	Postage (Reminder Letter)	5,540
1,278.75	Keypunching (IBM Service Bureau)	3,330
1,900.00	Analysis (ISR)	3,348
$ 6,897.96	TOTAL	

O.I.R. Salaries = Wages paid to two full-time employees for a twelve-month period

EXHIBIT A

THE UNIVERSITY OF MICHIGAN
ANN ARBOR

ALLAN F. SMITH
Vice-President
for Academic Affairs

March 18, 1966

Dear Colleague:

I never thought I would be asking your help by imposing a questionnaire upon you, but the job in the Office of Academic Affairs has clearly reached a point where I must do so.

Reliable information is essential if this office is to plan and decide upon priorities effectively. Your guidance is needed if the allocation of both time and money is to produce the most desirable improvement of the milieu in which you work. I know of no other way to get this information and therefore ask your cooperation and help. I have discussed the matter with the Educational Policies Subcommittee of SACUA and am reassured by the fact that they approve of my sending a questionnaire.

The attached questions are part of a survey to obtain a more accurate picture of how you and others feel about your work, your benefits, and conditions of work. The results will be used in trying to make the University a more stimulating and satisfying place in which to teach and do research.

I would emphasize that your answers will be kept anonymous. No attempt will be made to identify you. Your responses will be put together with others from your school, college, or other unit to provide us with a cross-section view of opinions on the various topics. Your questionnaire will be kept in this office for use only by those compiling the data. We will, of course, report the results to you as soon as they are ready.

The study depends upon the thoughtfulness and candor with which you answer the questions. Moreover, if it is to be successful, we must have a high rate of response. I hope each of you will take the 35-40 minutes to help us make better decisions.

Sincerely,

Allan F. Smith

Allan F. Smith

AFS/ji

Attachments

THEY WENT TO COLLEGE: A DESCRIPTIVE SUMMARY OF THE CLASS OF 1965

Robert J. Panos
Alexander W. Astin

American Council on Education

In the summer of 1965, the Office of Research of the American Council on Education followed up random samples of students from each of 246 colleges and universities included in a survey--conducted in the fall of 1961 at the National Merit Scholarship Corporation--of all entering freshmen students at a national sample of colleges and universities. These students would presumably, if they had proceeded through college at the normal rate (as many of them do not) have graduated in the summer of 1965. sample of 246 institutions was originally selected to be representative of the total population of regionally accredited four-year colleges and universities as of 1961 (Irwin, 1960). But because very small and relatively less affluent institutions would compose most of a simple random sample, the larger and wealthier institutions were deliberately over-represented in order to insure heterogeneity within the sample (Astin, 1965(a), pp. 102-105).

Questionnaires were mailed to a random sample of 60,078 of the 127,212 students included in the original survey--or approximately 250 students per institution. For institutions enrolling fewer than 300 freshmen in 1961, all students were included; random samples of 250 students were selected from the larger institutions. The arbitrary figure of 250 was chosen with the expectation that we would obtain about a 50 percent rate of return from the initial mailing. Following a reminder, 30,506 questionnaires were returned to the Council's Office of Research. A second, shorter form of the questionnaire was mailed to the 29,572 non-respondents, and an additional 5,899 forms were returned. The final number of usable completed questionnaires was 36,405--slightly more than 60 percent of the 60,078 students who were initially sent the follow-up questionnaire. However, since the U.S. Post Office could not forward the questionnaire to many of the students who had moved (the students' names and home addresses were obtained in 1961), we estimate that at least 75 percent of the students who actually received the mailed questionnaires eventually returned them.

The follow-up questionnaire inquired into several aspects of the student's progress: educational and vocational achievements, current activities, the college experience, current plans, and information about background not obtained in the 1961 survey. The purpose of this paper is to present this descriptive information about the Class of 1965, including data on the student's personal background and subsequent performance. Although such descriptive data do not establish causal relationships among the variables represented, we believe they are useful in placing the current vague and highly subjective notions about the outcomes of college into an objective, empirical frame of reference. Thus, this paper focuses on the descriptive or categorical rather than the inferential or theoretical aspects of the college experience.

Weighting the Data

Analyses were performed to determine the extent and possible effects of the bias in the sample of respondents. Although response biases did not appear to have any appreciable effect on cross-product relationships (that is, relative monotonic orderings among items), their obvious distorting effect on marginal tabulations makes it necessary to adjust the data to estimate population parameters. Thus we computed a set of individual weights to adjust for differences between respondents and non-respondents in the sample of 60,078 students to whom the questionnaires were sent, and also a set of institutional weights to adjust for the disproportionate sampling of institutions from the various stratification cells of the original sampling design (see Astin, 1965(a), pp. 102-105). The purpose of these weights was to adjust the data to be representative of the population of students entering accredited four-year colleges in 1961.

A multiple stepwise regression analysis was utilized to ascertain the degree of bias among respondents. This analysis was carried out by randomly selecting a sample of 4,000 students from the original mailing of 60,078 and defining for each student a dichotomous dependent variable (respondent or non-respondent) arbitrarily scored "0" or "1." A pool of 98 pre-college input items served as the independent variables in the analysis. The independent variables included sex, 12 high-level secondary school achievements (e.g., elected class president, placed first, second, or third in a school science contest), size of high school graduating class, overall high school grade average, level of educational aspiration, probable major field, initial career choice, father's education, and father's occupation. The .05 level was used as a cut-off for the stepwise procedure; in short, the final regression solution included those predictor variables whose independent contribution to the reduction of the residual sums of squares at each step was statistically significant at the .05 level of confidence.

The biasing variables in the sample of respondents were: the level of educational aspirations (less than a bachelor's degree vs. bachelor's or higher degree), high school grade average, level of father's education, and whether or not the student had ever had an original work published while attending high school. These variables defined a 7 x 4 x 2 x 2 four-way table into which the 36,405 respondents were sorted and individual weights for each of the 112 cells were computed. However, in order to obtain stable weights within all the cells of the tabulation, it was necessary to collapse two of the categories: level of aspiration and publication or nonpublication of original work.

The first results clearly show that high school grade average is monotonically related to response bias. Although father's educational level is also obviously related to the probability that a person will return a questionnaire, the relationship is not monotonic. Students whose fathers are high school dropouts are less likely than are other students, including those whose fathers have only a grammar school education, to return a questionnaire. The pattern of weights otherwise conform to logical, _a priori_ expectations that the least able students from relatively less educated families are not as likely to return a mailed questionnaire. These results indicate that it is essential to obtain some estimate of response bias in mail surveys and to adjust the data among the returns. A major obstacle to correctly interpreting and generalizing from the results of any mail survey is failure to take into account the degree of bias among respondents (Panos and Rice, 1967).

The marginal tabulations presented in this report are the result of applying the product of an individual's weight and his institutional weight to his data vector. Thus--to the extent possible given our sample and pool of items--the data reflect the results that would have been obtained if (a) everyone had responded to our mailed questionnaires (100 percent response rate), and if (b) our sample of institutions was a simple stratified random sample drawn from the defined population of regionally accredited four-year colleges and universities. The weighted descriptive data are based on the responses from 30,506 students who returned the initial long form of our questionnaire. The total weighted N was 622,413. Since everyone did not respond to every item in the questionnaire, the N varies slightly from item to item.

Background Information

Fifty-six percent of the students in the population were men; 44 percent were women. Slightly more than one-third of the Class of 1965 were married at the time of the follow-up survey. Of these, 5 percent had been married before starting college, over 58 percent while in college, and 36.3 percent after leaving college. In addition, 38.5 percent of this group had one or more children. The data support the notion that marriage is a more likely outcome of going to college for women than for men.

The students were asked to estimate their parents' current annual income and to indicate both parents' educational level. Only 7.3 percent of the Class of 1965 reported that their parents' incomes were less than $4,000; more than twice as many students said that their parents' incomes were over $19,000. The modal (21.4 percent) interval reported was $7,000-$9,999. Almost half the students' fathers had received at least some college training, and 42 percent of the mothers had continued their formal education beyond the high school level.

Slightly more than 96 percent of the students reported their racial background as Caucasian. Almost 3 percent indicated that they were Negro, 0.1 percent American Indian, 0.4 percent Oriental, and 0.3 percent "other." The data on racial background show that, in relation to the proportion of Negroes in the population (about 12 percent), the American Negro is under-represented among the college-trained-- particularly the Negro male: twice as many Negro girls as boys were members of the Class of 1965.

Over two-thirds of the students indicated that they came from Protestant family backgrounds, and 21.8 percent reported Roman Catholic backgrounds. Although only 2 percent checked "none" for their family religious background, 13.1 percent said that they themselves have no religious preference at present. These data suggest a trend among college students to reject the parents' religious beliefs, a tendency more pronounced among boys than among girls; a similar trend has been observed in other recent studies (Nichols and Astin, 1965; Astin, Panos, and Creager, 1966; Panos and Astin, 1967a). These data support the notion that one result of college, at least for some students, is a rejection of organized religion.

College Finances

Who pays for college expenses? How much money do college students earn from their summer jobs? How much money does the Class of 1965 expect to earn after their class has graduated? Items intended to shed some light on these questions were included in our follow-up questionnaire.

More than 75 percent of the Class of 1965 indicated that 9 of the 12 items listed were not a source of financial aid during their undergraduate years. As would be expected, parents are by far the largest single source of financial aid to the college student. Over 93 percent of the Class of 1965 reported that financial "support from their parents" helped pay for some part of their college and living expenses. The only other source of undergraduate financial support that a majority of students reported was their own earnings;over 87 percent of the students helped pay their own way. It would appear that the luxury of going to college--for the Class of 1965--was a financial burden borne largely by the parents and the students.

Slightly more than 11 percent of the Class of 1965 earned $1,000 or more during the summer of 1962 (at the end of their freshman year); more than twice as many reported earning of $1,000 or more during the summer of 1965. Over one-third of the students earned more than $600 in the summer of 1962, and 43.8 percent had earnings of at least $600 during the summer of 1965. As would be expected, the boys earn substantially more money than the girls; over half the boys earned $700 or more in 1965, whereas only 17.3 percent of the girls earned at least that much. However, the popular view that this earning power differential is attributable to discrimination against women in the employment arena should be tempered by the recognition that girls have a much easier financial go of it while in college than do the boys. Furthermore, parents with limited resources may be less likely to send their daughters to college. The notion that women who go to college come from relatively more affluent homes is supported by the data, which show that more boys than girls come from the lower end of the socio-economic scale. Thus, women may have less need for remunerative employment than do men.

The modal interval (34.4 percent) reported by the Class of 1965 as their anticipated 1966 average monthly incomes (before deductions) was $250-$499. As many as 5.1 percent expected to earn a monthly salary of $1,000 or more from their own work, and more than twice as many expected to be earning at least $1,000 a month from their total family income (not parental family) during 1966.

Undergraduate Employment Status

Students were asked also about the nature of the jobs they held both during the academic year and during the summer. Ten percent of the Class of 1965 were employed during the 1961-1962 academic year; of these, one-fifth reported that their jobs were career-related. During the 1965-1966 academic year (the time of the follow-up), three-fourths of the employed members of the Class of 1965 were doing career-related work. During the summer of 1962, over 60 percent of the students were employed. Of these, almost 20 percent reported that their jobs were career-related; however, during the summer of 1965, almost one-half of those employed were working in career-related jobs. These data indicate a trend over the college years toward vocationally oriented employment among the working members of the Class of 1965. The possible effects that the student's employment pattern during the college years has on his final career choice seems an area worth investigating.

Student Housing

Although 64 percent of the students lived in a college dormitory during their freshman year at college, only 24.5 percent were living in dormitories during the academic year 1964-65. As would be expected, the decrease over the college years in the percentage of students living in college housing is monotonic and apparently linear. The only other category of housing which shows a similar (but positive) monotonic trend over the years is the percentage of students living off campus in private apartments. Somewhat surprisingly, the percentage of students living at home with their parents increases immediately after the freshman year, but subsequently decreases. Women are more likely to live in a college dormitory throughout the college years than are men. The preceding observation lends support to the popular belief that in loco parentis--as applied by college officials--perpetuates our society's double standard with regard to differential treatment of the sexes.

Although more than twice as many boys as girls live off campus in a private apartment during the freshman year, more boys than girls live at home with their parents over the college years. Because the rules and regulations regarding student housing are one part of the college environment that can be manipulated rather easily by college and university officials, it would seem advisable that institutional researchers attempt to discover how these various housing arrangements affect the development of the student.

Follow-up Outcomes

The preceding sections of this report have discussed some personal and background characteristics of the Class of 1965. The following sections present descriptive data concerned with the students' activities and college experiences, their educational and vocational achievements, their attitudes toward college, and their plans for the future.

Because the data are based on student self-reports, the question naturally arises, to what extent do the results reflect inaccurate or deliberately distorted self-

The College: Academic Year 1964-1965

In this section, we present some of the typical activities and achievements of those members of the Class of 1965 who were enrolled as students during the academic year 1964-65. Over 21 percent of the students were members of a national or local social fraternity or sorority, and 25.4 percent were members of an honorary (subject matter) fraternity. Eleven percent were members of a college choir, glee club, or marching band, and 42 percent participated in extracurricular activities as members of college or intramural athletic teams.

Students were also asked about their experiences during the academic year 1964-65. More than one-third of the Class of 1965 reported that they fell in love during this period, and 19.8 percent said that they married. As many as 12 percent indicated they changed their major field of study in 1964-65. Almost 13 percent of the Class of 1965--more than twice as many boys as girls--reported that they had flunked a course in their senior year. One-fourth of the members of the Class of 1965 were elected to a student office, over 14 percent wrote an article for their school paper or magazine, 8 percent played on a varsity athletic team, and 2.7 percent had a lead in a college play.

The daily activities engaged in by 95 percent or more of both sexes were: asking questions in class, going to the movies, checking out a book or journal from the college library, discussing sex with friends, and studying in the library.

The College Image

The students who attended college during 1964-65 were asked to rate their institution with respect to their over-all satisfaction with it and with a number of specific aspects. Not surprisingly, the majority of students (83.7 percent) were satisfied with their college; slightly more than 10 percent said they were "on the fence," and only about 6 percent indicated that they were dissatisfied with their college.

Six percent of the students felt that there was too much social life on the campus, and 22.1 percent that there was not enough. Although 36.6 percent reported that they had not received enough advice and guidance from college faculty and staff, almost 80 percent said they had had just about the right amount of freedom in course selection. Three-fourths of the students said that the amount of work required in their courses was just about right; over 11 percent felt that the workload was not enough; and 13 percent that it was too much. Almost 40 percent of the Class of 1965 said that they did not have enough personal contact with faculty during the 1964-65 academic year, and 40.6 percent that there were not enough outlets for creative activities.

Over 50 percent of the students felt that there was much pressure to earn high grades, that most of their fellow students were of a high calibre academically, that there was keen competition among students for high grades, and that being in the college built poise and maturity. The pressure to get high grades was felt more by boys (54.7 percent) than by the girls (48.6 percent). Over one-third of the students said that they felt "lost" when they first came to the campus, and almost as many (30.8 percent) felt that most students on the campus were like "numbers in a book." Over two-thirds of the students said that the labels "social," "practical," "warm," "realistic," and "liberal" were very descriptive of their college. Only 18.5 percent thought that "intellectual" was very descriptive of the college atmosphere, and 15.5 percent said that it was not at all applicable.

The students who were in college during the academic year 1964-65 were asked to describe a course (taken during the year) that was most closely related to their primary field of interest. Over one-fourth of these courses had 14 or fewer students in the class, 55.7 percent had fewer than 25 students in the class, and 32.4 percent descriptions?

Our starting point is the assumption that the subject is willing to co-operate; that is, to be sincere and honest. The descriptive information presented here is based chiefly on responses to questions of fact rather than questions of opinion. Furthermore, the information requested is presumably known by the student and could, in theory, be objectively verified by an independent observer. Therefore, it is reasonable to suppose that an individual might deliberately give inaccurate information only if the question arouses conflict and defense. The nature of the items reported here, however, does not appear to invite faking--the items are not "content threatening." Thus, it seems safe to infer that the results presented here represent, for the most part, a reliable descriptive summary of the Class of 1965.

Persistence in College

How many get a terminal degree within the four years following their marticulation? How many students drop out of or transfer from college during the undergraduate years?

We estimate that 65 percent of the students in the population had completed four or more years of college at the time of the survey. Thus, over one-third of the Class of 1965 did not complete four academic years of college work within the four years following matriculation. Slightly less than 60 percent of the students had achieved a terminal degree; almost half the boys, though only 34.3 percent of the girls, held no degree at the time of the study. More than 44 percent of the students in the population had changed their college or dropped out of college since their matriculation in 1961.

Of the students who changed institutions or dropped out of college for any period of time since 1961, 16 percent reported they they were asked to leave their first college because of unsatisfactory academic work, over 80 percent indicated that they left voluntarily, and three percent admitted that they were forced to leave for disciplinary reasons. Almost two-thirds of the students who left their college of matriculation said that they would have left even if they had had greater financial resources at their disposal;slightly more than 63 percent of these students attended at least one other college since 1961. These data suggest that increased financial aid may not be the panacea, as some have suggested, for solving the problem of college student attrition.

That almost one-half of the students who left their first college indicated that they were dissatisfied with the environment suggests that, although great amounts of time and effort are expended annually by counselors, students, and parents in determining the "right" college, much more needs to be learned about the complex decision-process involved in selecting a college. Over 40 percent reported that they dropped out of college because they had changed their career plans or wanted time to reconsider their interests and vocational goals. Almost 30 percent of the girls said that marriage was a major reason for their decision to leave college, whereas only 7.8 percent of the boys reported marriage as a major factor in their decision.

The factors are, to be sure, the student's ex post facto explanations about why he dropped out of college, but they provide clues as to the types of items that should be included as input or control variables in longitudinal institutional studies of college student attrition. For example, items about the entering student's marital plans, his anxieties about college finances, and his degree of confidence in his expressed interests and career plans should be included in attrition studies to provide a frame of reference in which the later behavior of the drop-out can be viewed.

had 30 or more class members. Almost half the teachers held the academic rank of professor, and about one-third were assistant or associate professors. Although 18.4 percent of the instructors were women, only 5.6 percent of the male members of the Class of 1965 had a female instructor in the course which they described in this section of the questionnaire.

Educational Achievements and Aspirations

The modal (35.6 percent) over-all college grade average as reported by the Class of 1965 was B- or C+. Over 12 percent had a B+ or better over-all grade average, and 21.7 percent had a C grade average. The modal (34 percent) grade average in the major field of study was B. Less than 10 percent--and twice as many boys as girls--had a C over-all grade average in their major field of study.

In 1961, 57.6 percent of the Class of 1965 said that a bachelor's degree was the highest level degree that they planned to obtain. In 1965, only 23.2 percent reported the bachelor's degree as the highest academic degree they desired; 70 percent said they hoped to achieve a postgraduate degree, and 25.6 percent of these students hoped to reach the doctoral level.

Although 32.2 percent of the students acknowledged that they were not sure when they would obtain their highest degree, the data suggest that the Class of 1965 has somewhat unrealistic aspirations. A similar trend was reported in a recent study by Davis (1964), who found that more than three-fourths of a large sample of graduating college seniors expressed an interest in pursuing postgraduate training. If these data do imply such a trend in higher education, it would seem that the articulation between undergraduate and graduate institutions be even more carefully planned and evaluated than has been urged recently.

Discussion

The most readily available source of information about higher educational institutions is the student. Students are, in some respects, a captive audience: they have become accustomed to being asked to fill out a variety of questionnaires, forms, tests, inventories, booklets, and the like. The considerable interest that researchers and administrators have in student data (probably regarded by the students themselves as unnecessarily redundant) is easily understood. After all, most instructional processes are designed to produce changes in the student's intellectual and character development. Only by learning something about the student, and about how he changes during college, can those responsible for structuring particular learning experiences discover what their programs in fact accomplish. Information about the student is, in short, the core of any institutional research program.

Meaningful information about the student is vital because it extends the fund of knowledge necessary for rendering rational decisions. Administrative decision-making involves a choice among desired educational objectives and among the alternative means available for achieving these objectives. Therefore, it is important that the information consulted by the decision maker be relevant to the educational goals he considers desirable.

But the quality of available information can vary greatly. Furthermore, empirical data that are purely descriptive do not explore causal relationships among the variables represented, and it is the causal relationships that the administrator must know if he is to suit means to ends (Panos and Astin, 1967b). Although tabulations of descriptive data may imply certain trends, they provide no way of ascertaining the particular causal factors that account for the trends. For this reason, summary tabulations of data are useful if they are regarded as purely descriptive and as a possible source for ideas and hypotheses that can later be tested.

Summary

The purpose of this report was to make available a summary report of some of the results of a questionnaire survey, conducted in the summer of 1965, or 30,506 young adults who were enrolled as freshmen at a national sample of accredited four-year colleges and universities in the fall of 1961. The data presented were differentially weighted to represent the defined population of the Class of 1965. This report gave information concerning the backgrounds, personal characteristics, educational and vocational achievements, activities, and current plans of the Class of 1965. Certain implications of the data were discussed, but the primary purpose was to make these descriptive data available. It is our hope that with the current emphasis on institutional research, these data will be useful in suggesting ideas and hypotheses that can be tested in the context of such research.

References

Astin, Alexander W. Who Goes Where to College? Chicago: Science Research Associates, 1965(a).

Astin, Alexander W. The College Environment, Washington, D.C.: American Council on Education, 1967. (mimeo)

Astin, Alexander W., Panos, Robert J., and Creager, John A. "National Norms for Entering College Freshmen--Fall 1966," ACE Research Reports, Vol. 2, No. 1, American Council on Education, 1967.

Davis, James A. Great Aspirations. Chicago: Aldine Publishing Co., 1964.

Irwin, Mary. American Universities and Colleges. 8th ed. Washington, D.C.: American Council on Education, 1960.

Nichols, Robert C., and Astin, Alexander W. "Progress of the Merit Scholar: An Eight-year Follow-up," National Merit Scholarship Corporation Technical Report, Vol. 1, No. 5, 1965.

Panos, Robert J., and Astin, Alexander W. "A Profile of Entering 1965 College Freshmen," College and University, Vol. 42, Winter, 1967(a).

Panos, Robert J. and Astin, Alexander W "The Utilization of Information in the Educational Decision-making Process," Educational Record, American Council on Education, 1967(b). (in press)

Panos, Robert J. and Rice, Janet E. "An Experimental Analysis of Different Mail Follow-up Techniques Within a National Sample of Hard-Core Nonrespondents," to appear in Proceedings, American Psychological Association, 1967.

94/95

MEASURING SOME EFFECTS OF A "BASIC IDEAS" PROGRAM

Joseph C. Heston
Director, Bureau of Institutional Research and Counseling
Albion College

Abstract:

Albion College is experimenting with a "Basic Ideas" program for liberal arts freshmen and sophomores, an interdisciplinary sequence of four semesters. This report covers evaluation of the first year, comparing matched experimental and control groups. The Basic Ideas group shows a significant gain on the Inventory of Beliefs. Both groups made gains, but not significant, on the Critical Thinking Appraisal. The Basic Ideas group on the Educational Objectives list decreased vocational objectives and increased educational-cultural objectives. The rate of withdrawal from school by academically successful students was much greater for the Control group.

The Basic Ideas Program:

This is a four-semester sequence: Man and Self, Man and His Natural Environment, Man and His Social Environment, Man and His Values. It was planned and taught by a committee from the following disciplines: philosophy, religion, economics, psychology, biology, chemistry, mathematics, English, and speech. A primary goal was cultivation of the habit of mind to think critically about one's self, society, the natural order, and values. Critical awareness of problems and methods of approach was stressed, rather than a set of standard answers.

Class sections had ten students each to facilitate discussion and close interchange between faculty and students. Each teacher also served as the academic counselor for his section enrollees, a device for further personal contact. Fourteen different paperback books were used each semester as texts, roughly one per week. Paper writing and criticism was emphasized.

The Matched Groups:

Enrollees were chosen from volunteers. The whole freshman class was informed of the project by summer mailings. Two hundred fifty-eight of the 481 freshmen volunteered. Two matched groups of one hundred each were drawn up on the basis of sex, SAT Verbal scores, and general area of probable academic major or vocational objective. These groups purposely included students from all ability levels and the vocationally-undecided, as well as those with stated choices. Both the Experimental (Basic Ideas) and Control group were thus stratified, representative samples of the 258 volunteers. None of the "t" values show significant differences between the groups.

Evaluation Devices:

The following instruments were chosen to assess effects of the Basic Ideas program. The A.C.E. Inventory of Beliefs (1), the Watson-Glaser Critical Thinking Appraisal (4), and a modification of Lehmann and Dressel's Educational Objectives (2 and 3). They were to be given three times: (a) at time of freshman entrance, (b) start of sophomore year, and (c) end of sophomore year. At each administration the tests were given to all members of the Fall 1965 freshman class, so students were not aware there were experimental (Basic Ideas) and control groups.

This report covers the Fall 1965 testing and Fall 1966 (first) retesting. The data shown here are only those who took the tests both times; this includes eighty-eight of the Basic Ideas group, eighty-two of the Control group, and 247 Others. Hence, this is an interim report, pending the May 1967 (final) retesting. Changes evidenced on each of the three tests are considered separately below.

The Inventory of Beliefs:

The total score means and standard deviations for the Inventory of Beliefs are shown in Table 1. The Basic Ideas group made a fair gain, the Controls a slight gain, and Others a small loss. These changes, seen in terms of college freshman norms, are as follows:

Basic Ideas:	84 percentile to 88 percentile
Controls :	78 percentile to 79 percentile
Others :	76 percentile to 73 percentile

Variability decreased for the Basic Ideas group, but increased considerably for the other two groups. We had more low scores on retest, which may be partially attributed to a poor test-taking attitude seen in some of the sophomores.

Table 1. Basic Data For Inventory Of Beliefs

Group	N	Time Tested	Mean	S.D.	Test-Retest Correlation
Basic Ideas	88	Fall '65	74.330	15.102	.689
	88	Fall '66	77.330	13.528	
Control	82	Fall '65	71.024	12.087	.676
	82	Fall '66	71.512	15.165	
Others	281	Fall '65	69.330	13.300	----
	247	Fall '66	68.559	15.646	

The test manual reports test-retest scores for 2,546 students in thirteen colleges, tested in September and then again in May. Their initial mean score was 60.30 and the retest mean was 65.02. This gain is a bit more than ours, but our initial means were considerably higher. This makes later gains more difficult! For two samples the manual reports nine months test-retest correlations of .77 (N=77) and .73 (N=263). These are similar to our .689 for the Basic Ideas group and .676 for the Controls.

The Watson-Glaser Test:

Table 2 shows the Watson-Glaser total score means and standard deviations. All three groups showed some gain after one year of college. Two of the three, Basic Ideas and Others, showed increase in variability on the retest.

The mean changes, expressed in college freshman norms, are as follows:

Basic Ideas:	80 percentile to 84 percentile
Controls :	73 percentile to 80 percentile
Others :	64 percentile to 67 percentile

The test-retest correlations (.593 for Basic Ideas and .649 for Control) are considerably smaller than the odd-even split-half reliability coefficient of .85, reported for college freshmen in the test manual. Our drop in reliability is certainly due in part to the full year's lapse between testings.

Table 2. Basic Data for Watson-Glaser Test

Group	N	Time Tested	Mean	S.D.	Test-Retest Correlation
Basic Ideas	88	Fall '65	72.034	8.324	.593
	88	Fall '66	73.670	10.425	
Control	82	Fall '65	70.585	9.895	.649
	82	Fall '66	72.122	8.759	
Others	281	Fall '65	68.530	9.360	----
	247	Fall '66	69.340	10.469	

Table 3 breaks the Watson-Glaser data down by its five subtests for the Basic Ideas and Control groups. Both groups show small gains on the first four subtests, with slight drops on the fifth (Evaluation of Arguments). The test manual offers no norms for the separate subtests, since they are short and have rather low reliabilities. Our test-retest correlations are smaller than those quoted in the manual.

Table 3. Changes in Subtest Scores on Watson-Glaser Test

Group	Subtest	1965 Test	1966 Retest	Test-Retest Correlation
Basic Ideas (N=88)				
	1. Inference	11.48	12.24	.465
	2. Recog. Assumptions	12.48	12.82	.321
	3. Deductions	18.72	19.01	.507
	4. Interpretation	18.34	18.77	.368
	5. Eval. Arguments	11.02	10.93	.387
Control Group (N=82)				
	1. Inferences	11.44	11.85	.457
	2. Recog. Assumptions	11.88	12.55	.339
	3. Deduction	18.76	19.12	.599
	4. Interpretation	17.90	18.18	.488
	5. Eval. Arguments	10.61	10.41	.269

Significance of Observed Changes:

How significant are the test-retest changes shown in Tables 1, 2, and 3? We have computed in Table 4 the "t" value for the significance of all differences between groups and within groups. Table 4 shows more significant differences between groups than for retests within a group. These differences were also more significant in 1966 than in 1965. And to a large extent, the Inventory of Beliefs showed more significance between group differences than did the Watson-Glaser (W-G). The W-G total scores (used in sections (a) and (b)) show more dependable differences than do the W-G subtest scores (c).

The Basic Ideas group was significantly higher than either Controls or Others on the Inventory of Beliefs in 1965. The difference became more dependable (larger "t's") on the retest. The gain on Inventory of Beliefs within the Basic Ideas group from 1965 to 1966 was statistically significant at the .05 level and nearly large enough for the .01 level. We can be strongly sure the Basic Ideas group showed a real gain on the Inventory of Beliefs, i.e., they rejected more misconceptions after one year of college. Neither the Controls nor the Others gained enough on the Inventory of Beliefs to be statistically significant.

Table 4. Significance of Mean Differences

Groups	Time Tested	"t" for I.B.	"t" for W-G
a) Comparison between groups			
B. I. vs. Control	1965	1.933	0.796
B. I. vs. Control	1966	2.838**	1.350
B. I. vs. Others	1965	3.425**	2.698**
B. I. vs. Others	1966	5.300**	3.902**
b) Comparison within group (retests)			
Basic Ideas	1965vs.1966	2.473*	1.772
Control	1965vs.1966	0.390	1.766
Others	1965vs.1966	1.051	1.512

c) Comparison on W-G subtests

	(Test vs. Retest)	
W-G subtest	"t" for B.I.	"t" for Control
1.	2.569*	1.378
2.	1.240	1.957*
3.	0.785	1.090
4.	1.056	0.833
5.	0.743	0.674

"t" needed for significance:

* "t" > 1.96 for .05 level

** "t" > 2.58 for .01 level

On the W-G test the Basic Ideas group was *not* superior to the Control group to a statistically significant degree either in 1965 or in 1966, though "t" did increase in 1966. The Basic Ideas group *was* definitely superior to the Others on the W-G test both years, with "t" becoming larger in 1966. Section (b) shows none of the three groups made a statistically significant change on W-G when retested in 1966, although all did gain to some degree. The "t" values here suggest the observed differences might arise about 10 per cent of the time due to chance factors.

Section (b) in Table 4, with its disparity in "t's" for the three groups in Inventory of Beliefs test compared to the similarity of "t's" for the W-G test, suggests the Basic Ideas freshman year program did more to dispel misconceptions and erroneous beliefs than it did to increase reasoning and critical thinking! The W-G data indicate the total experience of one year in college increases critical thinking ability whether one takes the Basic Ideas program or not.

A supplementary analysis, not detailed here, was made of the Inventory of Beliefs and W-G changes for the Basic Ideas group at three levels on SAT Verbal--under 500, 500-599, and 600 and higher. The gains on both tests were greater at the higher verbal level, but the middle and lower verbal levels also showed definite gains.

Changes in Belief by Category Groupings:

In what areas of the Inventory of Beliefs does the Basic Ideas program seem to bring more change than in the control group? To avoid item-by-item details we arbitrarily grouped the 120 Inventory of Beliefs items into the eleven categories shown in Table 5. The number of items per category is shown in parentheses; they average eleven items per category. For each of our three groups we computed the average per cent of "correct" answers per each category in 1965 and again in 1966. For example, the Basic Ideas group in 1965 had 79.94 per cent correct answers in "Prejudices" items and 81.41 per cent in 1966, a gain of 1.47 per cent. This means they had fewer "prejudices" after a year in college.

Reference to the bottom of Table 5 shows the Basic Ideas group gained on more categories than it lost, five to two. The other two groups gained on fewer categories than they lost. Most of the gains or losses are not large, but since each is based on comparing two means, the differences are more stable than if one compared single item percentages.

Thus, the Basic Ideas group showed improvement in the areas of prejudice, culture and arts, sex and morality, God and religion, and family life. The Controls and Others likewise showed improvement in the areas of sex and morality, God and religion, and family life--but their gains were noticeably less than shown by the Basic Ideas group. Without testing statistical significance item by item, one cannot be certain of conclusions here. However, it seems safe to guess the larger Basic Ideas gains on these three categories indicate the Basic Ideas program has had greater impact here.

All three groups lost some on government and politics, business, and egocentrism. The last category may be interpreted that the retest found them more independent and feeling one's needs to serve himself.

Changes in Educational Objectives:

A detailed report on the characteristics and scoring of this Educational Objectives questionnaire was given at the 1966 AIR Forum (2). Therefore, the items are not reproduced here. The score used for each item was the frequency of the "No. 1" answers ("Very Important") given by each group. Test-retest correlations were .929 and .957 for the Basic Ideas and Control groups, showing very high consistency in the relative scores for each item from one year to the next.

Table 6 has been prepared to show what item changes might be significant. The letters and digits indicate the comparison made, e.g. B5:C5 means Basic Ideas 1965 vs. Control 1965. The chi square value is used to check the null hypothesis for each of the 120 comparisons. A chi square of 3.84 or larger is needed here to reject the assumption of no difference at the .05 level. For the significant differences, the comments which follow indicate the direction of the difference.

Table 5. Per Cent of "Correct" Answers to Various Categories on Inventory of Beliefs

Item Categories	No. of Items	BASIC IDEAS 1965	BASIC IDEAS 1966	BASIC IDEAS Diff.	CONTROL 1965	CONTROL 1966	CONTROL Diff.	OTHERS 1965	OTHERS 1966	OTHERS Diff.
Prejudices	(17)	79.94	81.41	+1.47	77.24	75.94	-1.30	72.24	70.53	-1.71
Cult. & Arts	(11)	69.90	72.09	+2.19	71.45	69.82	-1.63	69.36	70.64	+1.28
U.S.A. & World	(8)	70.37	70.50	+ .13	66.12	64.75	-1.37	61.00	58.75	-2.25
Gov't. & Polit.	(10)	63.50	61.60	-1.90	59.00	57.30	-1.70	62.10	56.80	-5.30
Sex & Morality	(13)	64.38	69.00	+4.62	59.08	61.62	+2.54	59.31	61.00	+1.69
Science	(8)	64.62	64.88	+ .26	59.62	60.88	+1.26	59.12	58.12	-1.00
Business	(11)	60.82	58.55	-2.27	57.91	56.73	-1.18	58.64	54.54	-4.10
God & Religion	(13)	62.15	73.38	+11.23	58.62	65.08	+6.46	55.23	61.85	+6.62
Miscellaneous	(11)	56.09	55.82	- .27	53.91	50.09	-3.82	51.45	48.36	-3.09
Family Life	(7)	44.00	51.43	+7.43	38.86	41.29	+2.43	39.57	41.86	+2.29
Egocentrism	(11)	39.54	38.91	- .63	37.36	36.45	- .91	34.00	32.09	-1.91
Total Average		62.64	64.82	+2.18	59.48	59.58	+ .10	57.70	57.12	- .58
	N	100	88		100	82		281	247	
Number of Gains over 1.00%				5			4			4
Number of Losses over 1.00%				2			6			7

Table 6. Chi Square Values -- Comparisons on "30 Objectives"
(Based on Number of "1" Answers per Item)

	Between Groups		Between Times	
Objective	B5:C5	B6:C6	B5:B6	C5:C6
1.	4.50*	0.12	1.48	1.56
2.	0.05	0.05	0.00	0.00
3.	0.18	0.02	0.09	0.00
4.	3.78	0.18	0.10	1.49
5.	2.43	0.51	2.76	0.40
6.	0.73	0.66	2.14	0.05
7.	0.04	4.89*	0.63	1.54
8.	0.57	0.40	0.11	1.09
9.	0.00	0.10	0.57	1.20
10.	0.51	0.00	0.02	0.63
11.	1.24	0.03	0.62	0.03
12.	0.32	2.56	1.12	4.15*
13.	0.00	0.71	0.10	0.23
14.	0.02	0.01	0.03	0.04
15.	5.39*	0.18	1.45	0.51
16.	0.03	0.44	0.02	0.39
17.	0.96	0.42	2.42	1.39
18.	1.20	0.00	0.61	0.11
19.	3.29	5.23*	0.00	0.22
20.	0.45	0.03	0.28	0.00
21.	0.94	0.02	2.81	0.64
22.	1.33	0.00	0.10	0.64
23.	0.47	0.04	1.03	2.12
24.	0.47	0.72	0.00	0.05
25.	0.60	2.30	7.98**	0.26
26.	0.00	1.16	2.35	0.14
27.	1.63	1.73	3.93*	0.40
28.	2.45	0.65	0.83	2.63
29.	0.39	0.05	1.53	0.65
30.	2.77	0.60	6.21*	2.38

For 1 d.f.:

P.10 = 2.71
* P.05 = 3.84
** P.01 = 6.64

In the first two columns, comparison between groups, there are four significant differences for these items:

1. To master classification of knowledge in a field.
15. To habitually apply scientific thought to discovery of facts.
7. To recognize the fact of world interdependence.
19. To understand one's physical and social environment.

Items 1 and 15 favor the Control 65 over Basic Ideas 65, indicating a more "applied" or "practical" reason to seek knowledge on the part of the Controls. Items 7 and 19 favor Basic Ideas 66 over Control 66, showing the Basic Ideas group gaining on Item 7 and increasing their "lead" over the Controls on Item 19.

Perhaps for this report, column three of Table 9 is most germane for the Basic Ideas group. This column shows for the freshman year a very significant change (decline) on Item 25 and significant gains on Items 27 and 30. These items are:

25. To develop ability to do significant independent research.
27. To develop a philosophy of life adequate for the twentieth century.
30. To become more critical of one's self, one's values, and one's world.

For the Control group only one item shows significant change over the year 1965-66. Item 12 shows a significant decline; it reads:

12. To understand other cultures and people.

Incidentally, the Basic Ideas group declined, too, but not significantly. Why should liberal arts education reduce this objective in both groups? Do they simply recognize our campus affords little direct opportunity for this?

Retention in School:

Is participation in the Basic Ideas program associated with a differential rate of retention in school? Seven of the Basic Ideas group did not return to campus for the sophomore year; fourteen of the Control group failed to return. Table 7 shows the reasons for withdrawal.

Since the Control group had so many "voluntary" withdrawals, it seemed useful to see what caliber of students were thus lost. Table 8 compares the voluntary withdrawals from each group with the returning members. The Control group lost many more students, and its attrition was among its best members, both in ability and academic success. Can we infer the Basic Ideas program was an influential factor in reducing voluntary withdrawals and avoiding similar attrition among its most capable students?

Table 7. Reasons for Withdrawal

Reason	Basic Ideas	Control
Voluntary	3	11
Academic grades	3	1
Medical	1	1
Personal	0	1
Total	7	14

Table 8. Comparison of Voluntary Withdrawals vs. Returning Members

Group	Mean Score	Voluntary Withdrawals	Returning Sophomores
Basic Ideas		(N=3)	(N=91)
	SAT Verbal	579	571
	SAT Math	653	588
	Freshman year GPA	2.65	2.70
Control		(N=11)	(N=82)
	SAT Verbal	578	563
	SAT Math	584	569
	Freshman year GPA	2.97	2.66

Student Opinion:

A very simple questionnaire was given the sophomore Basic Ideas students in December 1966. Their replies, shown in Table 9 show a strongly favorable attitude toward the program. We feel the consumers have endorsed our product!

Table 9. Questionnaire Replies from Basic Ideas Students, December 1966

(Entries are in per cent)

Question	Reply	Per cent
1. Has the Basic Ideas program contributed in a significant way to your education thus far?	Yes.	90.3
	No.	6.2
	Undecided.	3.5
2. Should the Basic Ideas program be continued?	Yes.	100.0
	No.	0.0
	Undecided.	0.0
3. If continued, should it be required of all students?	Yes.	14.3
	No.	84.8
	Undecided.	0.9
4. If continued, should it keep the same format?	Yes.	51.3
	No.	43.4
	Undecided.	5.3
5. Compared to other courses, how much time do you spend on Basic Ideas?	Significantly more	22.8
	Slightly more	45.7
	No difference	25.7
	Less	5.7

GPA Prediction:

As a supplement or by-product of this study, we examined these test data to determine their value in prediction of Grade Point Average (GPA) for the full freshman year at Albion College for all freshmen. Table 10 gives this information separately for 230 men and 230 women. Five separate predictors were used. The best of these were then tried in three possible combinations for multiple R's. High school rank, as might be expected, was the best single predictor. The Watson-Glaser test edged out SAT Verbal for second best predictor. The best combination of three predictors was high school rank, Watson-Glaser, and SAT Verbal.

Table 10. Prediction Data for Year I GPA

		Men (N=230)	Women (N=230)
a) First-order r's	Predictor		
HSR	High school rank	.426	.545
V	SAT Verbal	.350	.332
M	SAT Math	.297	.292
W-G	Watson-Glaser	.357	.423
I.B.	Inventory of Beliefs	.172	.070
b) Multiple r's	Predictor combination		
HCR, V, M		.533	.584
HCR, W-G, I.B.		.511	.609
HCR, W-G, V		.542	.612

References

Committee on Measurement and Evaluation. *Instructor's Manual for Inventory of Beliefs*. American Council on Education, 1953.

Heston, J. C. Educational objectives of liberal arts students, in *Research on academic input*. Association for Institutional Research, 1966.

Lehmann, I. J. and Dressel, P. L. *Critical thinking, attitudes, and values in higher education,* Michigan State University, 1962.

Watson, G. and Glaser, E. M. *Manual for Critical Thinking Appraisal*. Harcourt, Brace and World, 1964.

STUDENT ROLE ORIENTATION IN THE FRESHMAN YEAR: ITS STABILITY, CHANGE AND CORRELATES AT THIRTEEN SMALL COLLEGES

James V. McDowell, Research Associate
Project on Student Development

The Project on Student Development in Small Colleges is conducting a study at thirteen colleges of those students who entered as freshmen in the fall of 1965. Its purposes are to describe change in students as this occurs within a group of quite diverse colleges, and to study factors relevant to students' dropping out of or continuing in college.

During the first year of the Project, 1965-66, entering freshmen on all thirteen campuses were given a uniform battery of questionnaires and tests within the first week of their arrival. This procedure normally applied to the entire group of entering freshmen on a given campus, though there were some exceptions due to unforeseen events. The following spring a random sample of approximately 25 per cent of the same students was again administered most of the same test battery. Currently we are analyzing and reporting to the administrators and faculties of the participating colleges data from these two freshman test batteries. Also we are gathering further data this spring from a sample of the same students near the end of their second year of college.

The purpose of this report is to describe findings of the first year bearing upon student orientation toward college as this is assessed through a questionnaire technique based on the fourfold typology conceived by Clark and Trow (1966). Following a brief description of the technique, we shall consider these questions:

a) What degrees of stability and of change in role orientations of the students occur between the fall and the spring testing of identical individuals?

b) What cross-sectional relationships exist between role orientation data and data from other instruments, viz., the College and University Environment Scale (CUES) and the Omnibus Personality Inventory (OPI)?

c) As students modify their role orientations, in what ways are these changes related to data from the CUES and the OPI?

The Role Orientation Instrument

Clark and Trow (1966, p. 19) have described four "orientations toward a college education which are represented on American campuses and which may be in competition on any one campus." Identifying these as the Vocational, the Academic, the Collegiate, and the Nonconformist, they find them likely to flourish in a variety of mixed forms on any given campus and expect the individual student typically to participate in several of them, though usually finding his dominant orientation embodied in one.

The Clark-Trow typology questionnaire item consists of four paragraphs, each of which states a "philosophy of higher education" by describing what general goals are to be emphasized in college, what activities typically go with this view, what attitudes and involvements are typical of it. The student is asked to rank the statements in the order of their "closeness to your own philosophy of higher education" (Educational Testing Service, 1965). The student's "typology" is defined according to which of the four "philosophies" he selects as his first choice. In this questionnaire application of the concept, as Peterson (1965) notes, the intent is to describe not subcultures but the orientations of students toward higher education. It is concerned then, with students' motives and frameworks of thinking vis-a-vis the college environment in which they live. It is a broad characterization of the student stance, or role orientation, as the student himself reports it within the terms set by the item.

The principal research to the present on this instrument has been done by Peterson (1965). Through examining the relationship of typology responses to companion items in the College Student Questionnaire (Educational Testing Service, 1965), he established their substantial validity at the level of group data from a large sample of entering students at twenty-three American institutions of higher learning. Relationships consistent with the intent of the typology were found over a broad scope of topics, including intended major field, interest in extracurricular activities, preference for certain curricular and instructional policies, demographic factors, religious preference, attitudes toward parents and peers, and cultural sophistication. The instrument thus is one of wide relevance but of marked brevity, dealing with the student's motivational orientation to college.

In the present study, typology responses of 1988 students were obtained in the fall testing session, of whom 1147 were men and 841 were women. The spring sample made available responses of 460 students on the two occasions, 229 men and 231 women.

The distribution of responses from this sample as compared with preliminary comparative data on entering freshmen students at twenty-three American institutions of higher learning (Educational Testing Service, 1965b) appears in Table 1. The figures for Project subjects are exclusive of 10 per cent of the group, (not included in the figures above) whose fall responses were incomplete or otherwise defective.

Table 1. Percentage Distributions of Typology Responses

Group	Vocational	Academic	Collegiate	Non-conformist
ETS Sample	27	19	51	4
Project Sample, fall	26	26	40	8
Project Sample, spring	25	21	44	9

Preponderance of Collegiate orientations occurs in each group, somewhat less in the Project group than in the ETS sample. Nonconformists, the least frequent orientation in each group, occur twice as often in the Project group as in the ETS sample. The Project group showed a moderate shift toward Collegiate orientation between fall and spring, principally at the expense of the Academic Orientation.

Stability of Fall Typology Choice

To what extent do freshman students change their minds about first choice of "philosophy of education" between fall and spring? Comparison of the responses of identical individuals in the two testing sessions shows that slightly over half repeated the same choice in the spring, the others changing to some other first choice (Table 2). Fall choices in the Collegiate category were more likely to be retained than in other typologies, followed in order of stability by fall choices of Nonconformist. For choices of second, third and fourth ranking of orientations, fall-to-spring stability held in 45 per cent of the cases for men and 46 per cent for women.

Table 2. Per Cent Same Typology, Fall and Spring

	Vocational	Academic	Collegiate	Non-conformist	Total
Males	47	51	61	58	54
Females	35	30	71	42	51
Totals	41	41	66	50	53

Individual choices, in summation, show a moderate degree of stability through the freshman year, but also a substantial amount of shifting; and first choices show greater stability than the three later choices.

CUES and Role Orientation

The College and University Environment Scale (Pace, 1963) produces scores summarizing the student's perception of his college environment in terms of five variables, designated as Practicality, Community, Awareness, Propriety, and Scholarship. CUES scores for the Project colleges were available through responses of groups of 100 students on each campus, selected randomly within all four undergraduate classes, in the spring of 1966. Average CUES scores for each college were correlated with the per cent of students of each sex for each role orientation as expressed by entering freshmen in the fall (Table 3). The resulting coefficients seemed to run generally in patterns consistent with expectation, assuming that entering students had fairly accurate conceptions of the atmosphere of the respective institutions which they had chosen to attend. For example, the Vocational orientation, essentially practical rather than cultural, shows a negative relationship with Awareness -- a high score on the latter implying a high institutional value upon esthetic, philosophical and poetic interests.

Table 3. Correlations of Average CUES Scores
with Per Cent of Each Typology, Entering Freshmen*+

Typology (% 1st choice)		Practicality	Community	Awareness	Propriety	Scholarship
Vocational	M	.71	.04	-.58	.60	-.24
	F	.73	.23	-.59	.61	-.34
Academic	M	-.86	-.21	.50	-.64	.21
	F	-.71	-.25	.49	-.48	.36
Collegiate	M	.23	.42	.13	.36	.32
	F	.32	.56	-.08	.41	.05
Non-conformist	M	-.73	-.23	.44	-.62	.04
	F	-.62	-.32	.38	-.53	.06

---

*12 institutions for men, 13 for women.

+z-conversion values of r for samples of 12 cases take the following levels of significance (Guilford, 1950, p. 212):

r	p
.49	.10
.58	.05
.71	.01

A limitation of these data is the small number of observations involved. Simultaneously, the level of consistency between scores of males and females encourages confidence in a moderate level of stability of the coefficients for such a collection of colleges as this.

OPI and Role Orientation

Since a student's role orientation implies a stance having distinct motivational attitudinal qualities, one would reasonably expect systematic relationships between it and measures of personality bearing on cognate variables. Approaching this point empirically, the question was reasoned whether students expressing different role orientations would exhibit appreciably different patterns of scores on at least some of the fourteen scales of a broad scope personality instrument, the Omnibus Personality Inventory (Center for the Study of Higher Education, 1963). Average OPI scores of female students choosing each typology are presented in Graph I.

Appreciable differences in average personality scores for students of different typologies are clearly present. (Differences of two points are significant at the .05 confidence level among pairs of means not involving Non-conformist orientation. When Non-conformist means are involved, differences of three points are significant at the .05 level.) The trend of these differences may be seen in relation to the general content of OPI scales -- the four on the far left concerned essentially with intellectual processes; the next three with self-expression and religious orientation (high score indicating relative liberalism); the following to the right, with social relations and adjustment. For Practical Outlook, at the extreme right, high scores imply valuing material possessions, immediate utility, authoritarianism, and practical rather than theoretical interests. Non-conformists, as one extreme, thus tend to score high in intellectual, esthetic, impulsive, self-expressive, theoretical inclinations, but relatively low in social outreach and personal adjustment. Vocationalists, at the other extreme, follow precisely opposite trends on at least eleven of the fourteen variables -- though on some variables the distance between extremes is little. Academic and Collegiate subjects score at intermediate positions as a rule, and with notable consistency of ranking.

For male subjects (data not presented) the general pattern of results was the same, though the differences in scores among categories other than Non-conformist were less evenly spaced.

These observations may be compared with the summary sketches of the four typologies constructed by Peterson (1965) in the light of his study of College Student Questionnaire responses of students choosing the different typologies. In their bearing on student personality, his descriptive comments were these:

> Vocational Type: "His preference is to ingest passively rather than to explore or examine critically....He is apolitical and culturally plebeian."
>
> Academic Type: "...a genuine proclivity to explore, explain, and understand for the sake of understanding...his approach to learning is active, independent, and individualistic. He is skeptical of authority. He is introverted yet not disassociated. He is politically liberal and culturally aware."
>
> Collegiate Type: "The collegiate female is distinguished by her essential femininity by her overriding preoccupation with marriage and motherhood. For the collegiate type the important rewards are of an interpersonal nature...His characteristic approach to learning is at best passive and at worst chicane...He seldom is critical of authority (and)...is and aggressive extrovert. He prefers the forms of the popular culture over more serious idioms."
>
> Non-conformist Type: "(He)...is a seeker...(and) also a rejecter and a derider. Capable of genuine response to serious artistic statement, he is contemptuous of popular culture...(He) regards the classroom as yet another possible source of personally meaningful experience, ideas, and forms of expression

> ....In general he is alienated from the larger society, the 'system,' which considers him a misfit."

On the whole the profiles appear strongly consistent with Peterson's report, with the possible sole exception that Collegiate women are not appreciably outstanding in femininity (low Masculinity-Femininity score).

This finding of distinctive personality profiles for individuals of the various role orientations drawn from a multiple-college sample leads to the further question of whether role orientation may be a way of selecting individuals of reasonably homogeneous personality from different campuses. To explore this, colleges of distinctly different environment (as known from CUES profiles and a variety of other data available to the Project) were selected, the two criteria of selection being diversity of atmosphere and reasonably large numbers of students of a given role orientation and sex. Average scores of "Academic" men from three schools are presented in Table 4.

Table 4. OPI mean scores of Academic men from three selected colleges

	TI	TO	ES	CO	AU	RO	IE	SE	PI	AL	AM	MF	RB	PO
College A (n=17)	47	43	44	46	46	38	44	47	50	48	51	52	46	53
College B (n=32)	54	57	50	54	57	55	52	47	53	53	50	55	52	43
College C (n=45)	56	54	55	58	60	60	58	45	50	48	48	53	48	43

While there are variables in which differences among the three colleges are slight (e.g., SE, MF), there are more in which clearly significant differences occur (most dramatically, in RO--standard deviations average around 9 for the different variables and samples, and exceed 10 in only three of the 42 instances involved.) Data for the other typologies and for the other sex show fairly similar dispersions of OPI scores of students of similar typologies on campuses selected for contrast. The typology instrument thus is not a feasible means of selecting students homogeneous in personality from widely different student groups.

Correlates of Change of Typology

Typology Change vs. College and University Environment Scale.

If college environments have distinctive impacts on students, we would expect that role orientations might change through the freshman year in some systematic relationship to CUES scores of the colleges. This relationship was examined through the correlation of two sets of variables: the average CUES score of each college, and the per cent of increase or decrease from fall to spring in first choices of a given role orientation by identical groups of students on each campus. The coefficients obtained appear in Table 5.

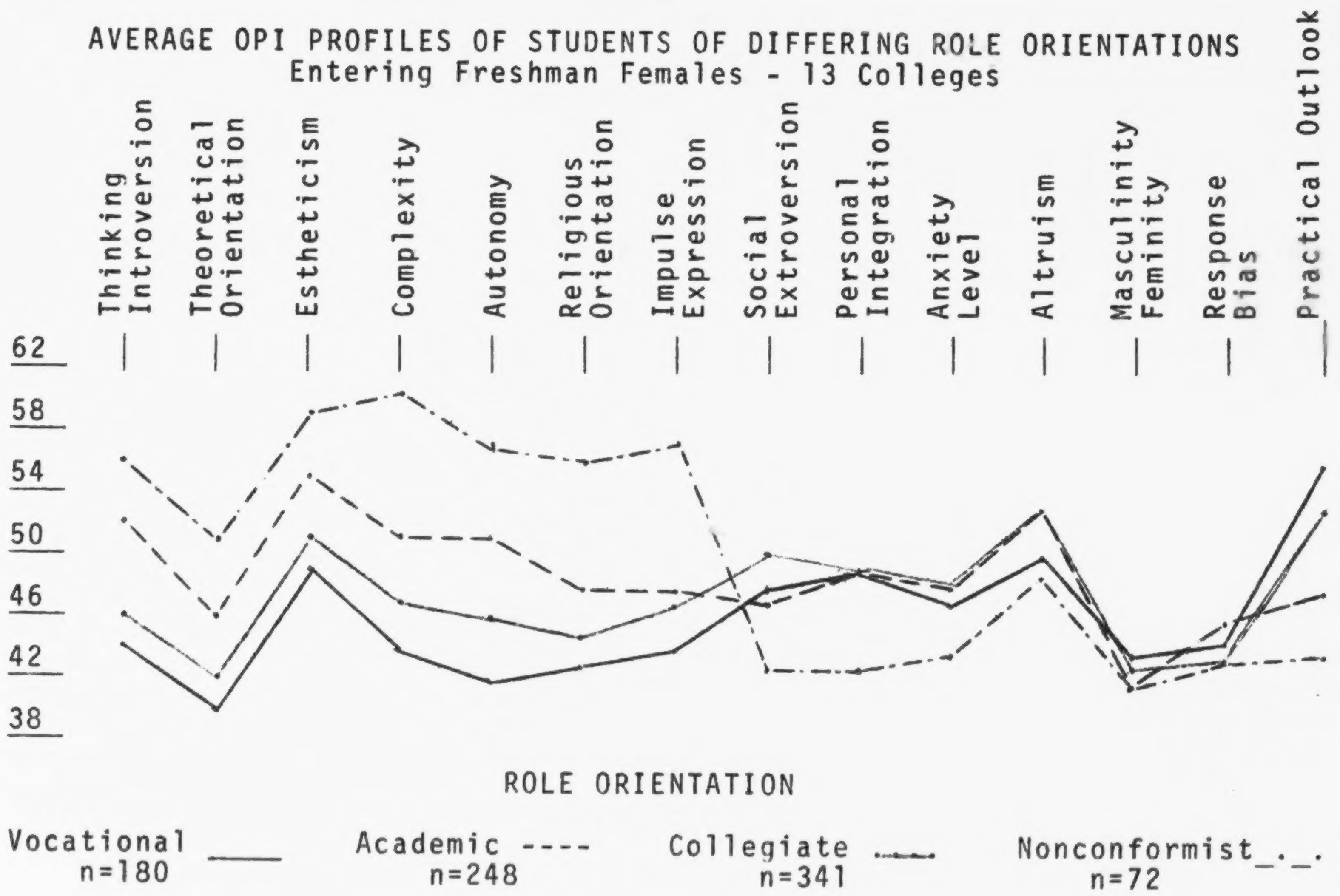
Table 5. Correlation Coefficients, per cent change in typology vs. CUES Average scores*

Typology		Practicality	Community	Awareness	Propriety	Scholarship
Vocational	M	.12	-.76	-.40	-.09	-.55
	F	-.03	-.60	-.34	-.21	-.44
Academic	M	.41	.22	.10	.02	.10
	F	.55	.10	-.36	.36	-.07
Collegiate	M	.04	.07	-.02	.40	.09
	F	.34	.26	.12	.55	.30
Non-conformist	M	-.49	-.38	.12	-.57	.02
	F	-.80	-.27	.44	-.61	.08

*12 Colleges for each sex.

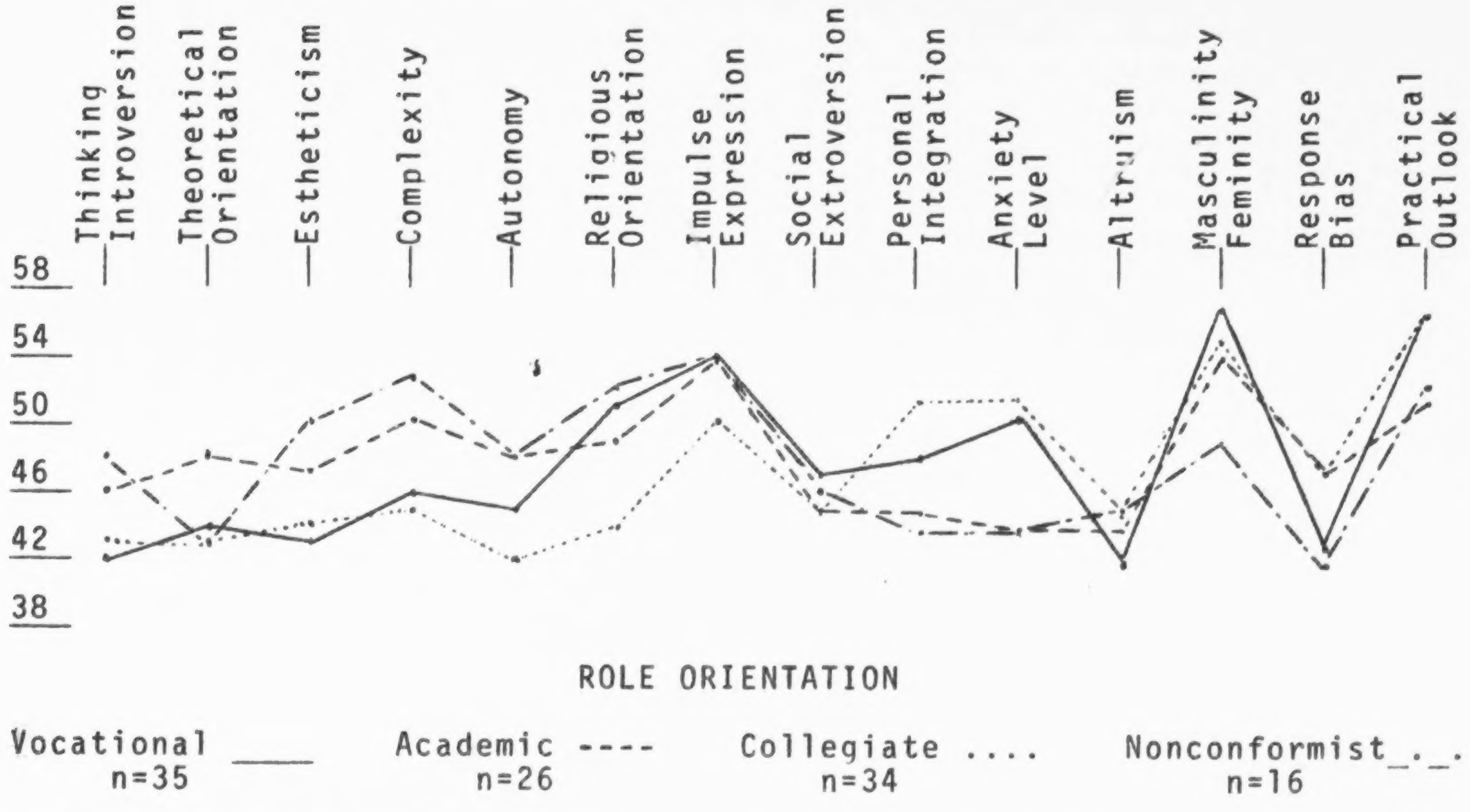
GRAPH I

AVERAGE OPI PROFILES OF STUDENTS OF DIFFERING ROLE ORIENTATIONS
Entering Freshman Females - 13 Colleges



GRAPH II

AVERAGE OPI PROFILES OF STUDENTS WHO LATER CHANGED TO SPECIFIED ROLE ORIENTATION
Entering Freshman Males - 12 Colleges



General consistency in the correlational patterns of the two sexes again is present. Some of the stronger relationships in evidence are:

CUES Scale	Role Orientation	Direction
Practicality	Non-conformist	Negative
Community	Vocational	Negative
Propriety	Non-conformist	Negative
Scholarship	Vocational	Negative

These relationships seem sufficient to justify more specifically focused research on the hypothesis that role orientations are influenced by the environmental forces suggested by the CUES scales involved. An attractive question, for example, is why a negative relationship occurs between the Community environmental influence and change toward Vocational orientation. Does a relative lack of warm social relationships on campus incline the student to focus more toward tangible future goals? And will a similar relationship hold for the upper-class years, when warm relationships on a co-ed campus might tend to encourage vocationally-oriented thinking?

Typology Change vs. Omnibus Personality Inventory

As indicated above, there is a tangible consistency between personality and role orientation at the time of entering college. Hence, it would be reasonable to assume that those students who changed to any given orientation did so in part because the orientation to which they changed was, in their several colleges, temperamentally more congenial to them than the one with which they entered. In such case we would expect that the fall OPI scores of those individuals who later changed to a given orientation would be distinguishable from those of individuals who changed to other orientations -- and that each group of changers would tend to resemble, in OPI patterns, the group of those who expressed a corresponding choice of orientation at the time of entering college.

Graph II displays the August fall OPI profiles of students who changed first choices of orientation in the spring to the ones indicated. Similarities to the basic fall profiles are readily observable. The same tendency for Non-conformist to range high toward the left of the profile, then to become lowest, is again apparent, while the other types show considerable consistency with profiles of the corresponding fall typologies. A number of significant differences among means are present, the difference required for significant at the .05 confidence level ranging from four to seven points and averaging about five points. The corresponding female data are roughly similar, though showing greater divergence of Non-conformism from the other orientations and less divergence among the latter three.

Conclusion

The findings of this study add to earlier evidence on the usefulness of the role-orientation questionnaire item as a means of studying college student groups. Typology responses show coherent relationships with CUES and OPI in cross-sectional data, in some cases easily according with expectation, in other cases calling attention to relationships meriting more intensive probing. Even with rather small numbers of observations, they are sensitive enough to show statistically reliable relationships between their changes during the first year of college and measures of environmental and personality variables. The small time required for administering the role orientation item and the relative ease of communicating data from it are additional advantages.

References

Chickering, Arthur W. "Institutional Differences and Student Characteristics," Journal of the American College Health Association, (December 1966), 168-181.

Center for the Study of Higher Education. Attitude Inventory: OPI Form D. Berkeley, 1963.

Clark, Burton R. and Trow, Martin. "The Organizational Context," T. M. Newcomb and E. K. Wilson, eds., College Peer Groups. Chicago, 1966.

Educational Testing Service. College Student Questionnaire, Part 2, Form 200D. Princeton, 1965a.

_______. Preliminary Comparative Data for the College Student Questionnaires. Princeton, 1965a.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York, 1950.

Pace, C. Robert. Technical Manual, College and University Environment Scales. Princeton, 1963.

Peterson, Richard E. On a Typology of College Students. Princeton, 1965.

THE INVOLVEMENT OF CURRICULA IN STUDENT COURSE CHANGES

Bernard B. Hoffman
Associate Director, Institutional Research
Syracuse University

At the beginning of each college term in the United States, over a quarter of the student body change one or more courses within a week or ten days following registration. This paper is a critique of the findings of a multi-institutional study which explored the reasons for this phenomena and in particular, the involvement of curricula. In accomplishing this task, we touched some ninety other colleges and universities with enrollments of over one-sixth of the nation's college students. The magnitude of this problem, which in some institutions is referred to as the "second registration," can be given the following dimensions. In one western college it was found that over 70 per cent of the student body was involved in changing some 25,000 course enrollments last spring. In another institution in New York State, less than 1 per cent of the student body was involved. This paper will explore the following areas: (1) why students change courses and (2) why "student originally selected course" later dropped.

Before beginning our discussion, a brief look at our sample and method is desirable. This study began with a pilot study at Syracuse University in the fall of 1965. At this time, we tested our questions and attempted to identify areas for future research. Perhaps the single most important discovery in Phase I of the study was the identification of the drop as the triggering behavior for over 90 per cent of the courses added. As a result of this finding, Phase II and Phase III of the national study were oriented toward dropping behavior. A brief look at the questionnaires used in these studies indicates that they ask the expected questions: what class standing, sex, college and major, full-time and part-time status, time of registration, alternative course selections, the department, course and major area in which course is being dropped, first time or repeated experience in schedule changing, and last, two unstructured response questions: why "student originally selected course" dropped, and why the student is dropping the course. It was these two latter questions which provided the most rewarding data and, in effect, answered the question, "Why Do Students Drop Courses?".

The sixteen categories which were identified by the unstructured responses were later combined into a twelve-category, forced-selection questionnaire with one escape, "other reasons," which was used in the final phase of our national study. The accuracy of the categories evolving from the unstructured responses is indicated by the fact that 87.8 per cent of the students who responded from some ninety institutions found one of the twelve categories to reflect satisfactorily their reasons for dropping a course.

A brief look at our sample taken in terms of population reveals: 30 per cent was drawn from the East and Middle-Atlantic, 33 per cent from the Central States, 18 per cent from the South, and 19 per cent from the West and Far-West. The number of students enrolled in the participating institutions at the last tabulation was 938,000. Over 5,000 samples were taken from this universe over the two-year period of this study. (Smith and Hoffman's "Registration Roulette" provides detailed related material and has one section devoted to the local application of techniques of identification, measurement and evaluation of schedule-changing problems on the local campus level, which should prove of considerable value to the university or college facing a sizeable problem in this area.)

Why do students drop courses? Actually, the question should be not only "why" do students drop courses, but "who" are these students who drop courses. To answer the latter question, we determined on a per capita basis, with a sample of 1848, that it was the freshmen who accounted for approximately 29 per cent, sophomores 24 per cent, juniors and seniors about equally at 21 per cent, and the graduate students about 4 per cent of the adding and dropping phenomena. Comparing graduates and undergraduates, the per capita rate of changing indicated that the undergraduate had roughly twice that of the graduate student. In terms of course enrollment, (at Syracuse University only), this raised the graduates' portion by approximately 20 per cent but left the undergraduate dominance of schedule-changing clearly evident. In terms of per capita changing on the basis of sex, we found that the female enrollment of the institutions sampled was 36 per cent. However, they accounted for 41 per cent of the sample which indicates a somewhat higher schedule-changing rate by our female students.

In terms of rank order, the reasons for dropping courses were as follows: Curriculum load too great, 12.6 per cent; Misinformed (source unidentified), 12.2 per cent; Seeking an easier course (Mickey Mouse type), 12.0 per cent; Mechanical conflict (hours and days), 11.5 per cent; Academic difficulty, 9.5 per cent; Poor Instructor or course, 7.4 per cent; Catalogue description, 6.8 per cent; Non-academic difficulties, 6.3 per cent; No response, 5.2 per cent; Questionable credit value, 3.6 per cent; Section changes, 3.5 per cent; College or Major requirements, 3.0 per cent; Will take course later, 2.4 per cent; Instructor advised drop, 1.8 per cent; Student changed major, 1.3 per cent; and Misadvised, 0.9 per cent. These were the categories resulting in the tabulation of the unstructured responses, and under each one, there are several sub-categories which would put the category in much sharper perspective. Basically, these are the responses of the students in their own words which account for the inclusion of the unresponsive answer, "Will take course later," which occurred 2.4 per cent of the time.

In summary, freshmen and graduate students were found to dominate the responses of Load too great and Academic difficulty. In this latter response, over 22 per cent of all freshmen and 20 per cent of all graduates responded that they were in academic difficulty, and this was the reason for their schedule-changing. The seniors at our colleges and universities appear most frequently the victims of misinformation and omissions in the catalogue. In addition, they are most frequently found changing schedules due to mechanical conflicts, due primarily to the hours and days of course offerings. Over 55 per cent of all graduate student responses identified the poor instructor or course as the reason for their schedule-changing behavior. In addition, upper-division undergraduates in their major areas frequently gave a similar response. Male students, in general, dominated the categories of Academic difficulty, Non-academic difficulty (referring primarily to part-time employment), and No response, accounting for over 70 per cent in each of these groupings. The female students were found most frequently related to areas of Misinformation, of Catalogue description, of Misadvising, and Instructor Advised drops. Furthermore, the female student dominated the categories of Changes in Majors and Section Changes. Courses dropped in non-major areas, as we would expect, are found highly related to Academic Difficulty. Courses dropped in students' major areas, as we have already indicated, were associated with Poor instructor or course and Instructor advised drop.

The reasons for selecting courses later dropped were as follows: Requirements, 44.3 per cent; No response, 19.2 per cent; Enjoyment, curiosity or self-improvement, 10.2 per cent; Catalogue description, 8.7 per cent; Part of a sequence, 8.2 per cent; Advisor's suggestion, 3.8 per cent; Mickey Mouse characteristics, 3.8 per cent; Hearsay, 2.1 per cent. We observe that the required courses were most frequently associated with the freshmen and graduate students, and male students in general. The reasons for dropping required courses were dominated by Changes in majors, as expected, Questionable credit value, Misinformation, Load too great, and Academic *difficulty*. Courses selected for *Personal Enjoyment, Self-Improvement* were most characteristic of the senior year and least among freshmen and graduate students. Reasons given for dropping these courses were to *Find an easier course* because of *Poor Instructor*, *Part-time employment*, and *College or Major requirements*. Courses selected because of their *Catalogue description* were most frequently made by seniors and female students and were later dropped to *Seek easier courses*, or due to *Poor Instructor or course*. Courses in a sequence were found to be dropped because of *Misinformation* and *Academic difficulty*. It was discovered that courses selected at the *Advisor's suggestion* were frequently dropped due to *Errors in the Catalogue*, which indicates both student and advisor had been misled. Courses selected due to their Mickey Mouse characteristics were most typical of the senior and graduate student and occurred twice as frequently among the male students. The primary reason for dropping Mickey Mouse courses was attributed to *Misinformation* and *Instructor advised*. As a sidelight, the *Instructor advised* drops were found to be most associated with these students seeking Mickey Mouse courses, but also associated with the student who took a course out of curiosity, enjoyment, or self-improvement. Finally, selections which were based on hearsay from fellow students were found to be most characteristic of the senior year and most frequently dropped for a number of reasons: *Load too great*, *Poor Instructor*, *College or Major requirements*, *Will Take course later*, or *Instructor advised*.

Summarizing our findings and drawing our data from the departmental and course analysis, we are able to generalize broadly why students drop courses. There were definite indications of curriculum involvement by way of required courses and courses in sequence which were producing a disproportionate level of academic difficulty. In addition, there were indications that some subjects were not adequately matched to the high-school achievement of freshman students, and thus producing academic problems. These are serious matters involving the student's adjustment and survival which should be rectified in the counseling and testing area prior to registration. At Syracuse University, an effort was made in this direction by the Romance Language Department, specifically Spanish, resulting in a reduction of over 60 per cent of the schedule-changing. This is just one example of what can be done.

Another generalization involves mechanical conflicts, which is a major reason for student schedule-changing. The departmental involvement clearly indicates considerable variation in the performance of scheduling in the various departments. This matter can become so serious, in fact, that it can invalidate a priority system of registration as a result of an overflow of course offerings during the prime hours of the day and week.

A final area of generalization involves the area of communication about our curriculum. We discover that catalogue selection was clearly linked to catalogue and advising errors. Hearsay from fellow students frequently proved unreliable. Data not appearing in our catalogue, of course, about poor instructors and poor course offerings, were a major cause of dropping behavior among graduate students and undergraduates in their major areas.

Finally, we wish to note once more the natural tendencies of registrars and scheduling officers to bewail the adds and drops and blame the student who often appears to be playing a game of "musical chairs," seeking Mickey Mouse courses, and avoiding normal levels of academic performance. This report recognizes that there are many such students on our campuses, but they appear to be a small minority of about 12 per cent. This minority suggests that much of the answer as to why students change courses represents an honest disappointment in our means of communication about curriculum, the matching of student achievement to course level, the mechanics of scheduling, or the more serious alienation from instructors and the curriculum itself.

DIFFERENTIAL FAILURE RATES IN COLLEGE: IMPLICIT EDUCATIONAL CONCEPTS
(A preliminary analysis)

Dr. Charles Bahn, Acting Director
Miss Mary Cox - Mrs. Dorothy Lee
Office of Institutional Research Services
The City College of the City University of New York

The Question of the Meaning of Grades

Ideally, a grade has several functions. It is an evaluation of a student's achievement and performance. It is a stimulus and a guide to learning. And it is a criterion on which prediction of the future performance of a student may be based. The factors affecting a grade evaluation are numerous and interdependent. The student factor includes his ability, preparation, knowledge, and interest. The instructor sets the basis on which a grade will be based, whether it is mastery of a skill or subject matter, development of sensitivity and appreciation, or critical judgment. The instructor also decides upon the methods of evaluation and determines his particular marking system. He may have absolute standards, relative standards which carry over from one year to the next, relative standards within each class; or he may gauge a student's achievement relative to his own ability, that is, measure his progress in the course.

Considering the great variability possible in grading, the question of the "meaning" of a grade arises. Does a particular letter grade mean the same from different instructors or from different departments or schools? Is a grade average a valid indicator of student achievement?

Our study of grade distributions, analyzed from the point of view of failure rates, indicates great variability from department to department, from course to course, and even from instructor to instructor. Hence, how is the student, the institution, or our grade-conscious society to interpret the omnipresent GPA? Why, given the variability of meaning of grades, does the grade point average remain one of the best of the several poor predictors of academic performance in use?

Failure Rates

This analysis takes as its unit the failure rate in courses. The term failure as used in this report includes grades of D (merely passing), E (conditional failure), F (failure), and where recorded, G and H (both, drop with failure). "D" is included because, while not a failing grade, "D" is assigned the value of minus-one in the computation of the student's average, and a specified number of D's in the record of a "C"-average student will result in his expulsion. The report includes comparisons of failure rates in different subject areas, in different levels (specifically basic, including introductory and prerequisite courses, vs. advanced and elective courses), and in different sections of the same courses.

All grades reported are for the undergraduate-day session of The City College for the time periods indicated. Data are reported separately for the Baruch School of Business and Public Administration of The City College of New York and the Uptown Campus, which includes the College of Liberal Arts, the School of Engineering and Architecture, and the School of Education.

Subject Areas

During the period of Fall 1962 through Fall 1965, the average failure rate for all courses on the Uptown Campus was roughly 12-13 per cent, while that of the Baruch School was 15-20 per cent.

In all courses taken at the Baruch School during the Fall of 1965, 44 per cent of the grades in mathematics courses were failures, as were about one-third of the grades in the Romance languages, one-fourth in history, and one-fifth in English. The failure rate in the physical sciences was less than one-fifth; in business courses, a range of 8-19 per cent.

On the Uptown Campus during this period, the mathematics courses again had the highest proportion of failures (about one-third). Physics was next, with slightly less than one-quarter failing, and in French and German 23 per cent and 22 per cent failed, respectively. In chemistry and civil engineering the rates were 22 per cent and 21 per cent respectively, while in all other subject areas fewer than 20 per cent of the grades were failures.

Considering non-professional courses alone (i.e., excluding engineering, architecture, education, etc.), the mean percentages of failure on the Uptown Campus for the years 1962-1965 show that 30.27 per cent of math grades were failures, with 24.11 per cent in physics, 23.95 per cent in chemistry, and 19.37 per cent in French. Basic biology failure rate was 18.52 per cent, Spanish 13.47 per cent, economics 12.75 per cent, social studies 11.65 per cent, and philosophy 10.7 per cent. In other subject areas failures were less than 10 per cent of the total grades -- history 8.86 per cent,political science 8.7 per cent, English 8.07 per cent, and psychology 5.2 per cent.

Proportional Failure Related to Proportional Grade Contribution

Some of the major disciplines contribute a greater proportion of failures than of grades to the total Uptown figures. Mathematics and the physical science courses contribute one-fifth of the grades, but about two-fifths of the failures. Mathematics, the physical sciences, and engineering contribute over one-quarter of the grades, but about half of the failures. If Romance language courses are added, the group contributes one-third of the grades, but almost two-thirds of the failures.

There is much year-to-year variability in the failure rates of courses, though failure rates averaged for subject areas are consistent over time.

In all subject areas offering basic courses, these courses contribute a disproportionately large percentage of failure grades. This difference is statistically significant. For those subject areas with more than 200 failing grades (history, English, and all the physical sciences), over 60 per cent of failures are in basic courses. This basic course failure syndrome appears to be not just a fact of our academic life, but also a growing trend. Over the past four years, increasingly higher proportions of failing grades have occurred in basic courses.

Multi-Section Courses

When per cents of failures in the sections of a single course are compared, great variability again appears. We derived data from grade distributions for courses taught during Fall 1966 semester on Uptown Campus. The sections of a course generally include similar numbers of students, as sections are closed when enrollment reaches a certain maximum. Further, there are no selection criteria other than student preference to systematically affect assignment to sections. Thus, grade patterns could be expected to be comparable; however, taking the difference between the largest and smallest failure rates in similar-size sections of a single course, one finds a maximum discrepancy of 61.07 per cent.

Discussion

The variability in grading patterns described above could create some interesting problems for the student aware of the significance of grades in his academic and subsequent career. For example, during the year 1965-66, a mathematics major, working towards a B.S., had to take courses where the failure rate was at the 30 per cent level, while a psychology major seeking a B.S. degree faced courses with only a 5 per cent rate (in courses in his major).

The B.S. student was also confronted with a difference between failure rates in the B.S. requisites and failure rates in the B.A. requisites. Finally, the B.S. student is neither required nor urged to sample the liberal arts courses beyond three basic courses. In any choice in a non-basic liberal arts course, he would have risked, for example, only a 7.7 per cent (political science), 6.1 per cent (English), or a 4.85 per cent (history), chance of failure.

The lot of the beginning B.S. student is increasingly hazardous as evidenced in the trends of the failure rates in the past four years in basic courses. The basic course experience often determines the student's further involvement in a specific area. Basic course failures in mathematics are up from 17 per cent of all basic course failures in 1963, to 25 per cent in 1966; in physics from 9 per cent to 11 per cent and in Fall 1966 prospective engineers faced courses ranked consistently high in failure rates, illustrating that the prospective engineer places himself in a continued "high risk" academic situation by his choice of a major.

Faculty Interviews

To explore the role of the individual instructor in determining the meaning of grades, thirty-two faculty members in nine departments of our Uptown Campus were interviewed about their opinions and attitudes on grading. Instructors who teach or have taught both basic and advanced courses were selected. The interviews were of approximately thirty minutes duration and were semi-structured, using a general questionnaire as a guide or point of departure. Topics included in all interviews were the function of grades, the method of grading, intra-departmental discussion on grading, personal evaluation of the present practices and system, and suggestions for changes in the system. Our description and analysis of the interviews is organized into these topic areas.

It must be noted that the interviews were based, to a large extent, on the spontaneous comments of those interviewed, rather than on a detailed survey on each topic. Thus, the report of a specific percentage of respondents supporting a particular view does not suggest that the remaining necessarily fail to concur. Rather, this view was not among the most salient for them, and was therefore not discussed.

Our sample of the faculty regards grades as serving many functions. The use of college grades as a criterion of admission by graduate and professional schools is considered a major justification for their existence. Fifty per cent of the faculty interviewed referred to this aspect of the grading system (all percentages based on 32 = 100 per cent).

A large portion of those interviewed (ca. 50 per cent) considered grades an evaluation of a student's performance, which may be helpful to the student in giving him an indication of the standards he is to meet and in offering a critical response to his work. All respondents in the mathematics and physics department mentioned this role of grades as a fairly accurate, evaluative guide for the student. All the English instructors interviewed also thought that the grade is a valuable critical feedback to the student. (Basic English courses have a conference system, in which each student meets individually with the instructor several times during the semester). Forty-four per cent of the instructors regarded grades as a simple record of performance, for use by the college (or by the mysterious black box, referred to as "the system"). The motivational aspects of a grade were also cited by a large number of faculty (ca. 42 per cent). In this respect, grades are seen as an incentive to the student, with emphasis on either the reward or the punishment aspect. Mention was made of the many grade-oriented students who push an instructor into using grades as a motivational tool. Among other functions of grading, future employment, an evaluation of instructor performance, and weeding out of unqualified students were mentioned.

Faculty members differed considerably in their methods of grading and in the selection of criteria on which they base their grades The criteria used, in varying combinations, are examinations, papers, laboratory work, classroom participation, and subjective impressions. Most instructors grade on an intuitive curve, which shows some variation from course to course and year to year, yet retains a certain inherent stability. Very few instructors grade on an explicit curve. In the physical sciences, mathematics, and Romance languages, there are more examinations during the semester, and the examination average is the major factor in determining the grade. A subjective evaluation of classroom or laboratory performance is only a very small factor in grading and becomes important only in borderline cases. (Borderline between two consecutive grades) In the advanced courses in these three subject areas, papers, laboratory work, and class participation become a larger component of the grade, but examinations are still of importance in most courses. In the social sciences and humanities (excluding Romance languages) instructors place a larger emphasis on more subjective criteria, such as papers, class discussion, and an overall impression of the student. Examinations also tend to be more subjective, frequently including more essay questions.

In all departments a difference in grading practices between basic and advanced courses was described. Some faculty explicitly cited these differences, others implied them. In basic courses there are, usually, several sections with different instructors. The majority of departments administer departmental finals in these courses. Several instructors were concerned with the great variation in grading distributions between instructors for the same course, particularly evident in these basic courses. In an attempt to achieve greater uniformity, one department has instituted a committee-marking system, in which one instructor marks one part of the examination for all students in the large multi-section introductory course.

In advanced courses most faculty members indicated that they expected work of higher quality from their students, and often there is an increase in the quantity of work demanded. However, there seemed to be general agreement that grades were nonetheless higher in advanced courses. Among the reasons given for this change are a greater student interest (often a large percentage are majors in the given department), an increased willingness to work, a weeding out of poor students in the basic courses, and a pressure for good grades for graduate school.

Discussion about grading practices in virtually non-existent on an interdepartmental level. In the social sciences and humanities there is even relatively little discussion about grading within the department. However, the faculties of the physical sciences and mathematics departments do discuss grading practices, partially prompted (in the physical sciences) by a necessary agreement between the lecturer and the laboratory instructor on each student's grade.

More than half of the faculty members interviewed expressed dissatisfaction with their grading practices and with the system as a whole. Of these, many considered grades an unreliable, and, at times, even arbitrary indicators of performance. Several instructors expressed concern with the absence of any absolute standards. The and lack of understanding of grades by both students and the "system" was another source of frustration. Rather than being a reward or the punitive tool of a disciplinarian, grades should serve a communicative function between teacher and student, and, as such, would be of value in the educational process. Several faculty members felt that grades would be unnecessary if it were not for the "system." An oral or written evaluation would serve the same educational function, and would eliminate the "evils" of grades. Some instructors see grades as actually inhibiting learning. Concern was expressed by a number of faculty members about the great responsibility for the student's future placed on the individual instructor. The present draft system and its relation to the war in Vietnam has increased this responsibility with its attendant pressures and discomforts. The large number of grade-oriented students is also disturbing to several faculty members. Twenty-seven per cent of the respondents in the departments of physical sciences, mathematics and Romance languages expressed dissatisfaction, whereas 82 per cent of respondents in other departments indicated such dissatisfaction (departments of economics, English, history, political science, psychology).

Almost all the faculty members recognized some need for a change in the present system, (81.25 per cent). However, proposals for change are quite varied, ranging from a refinement of the present grade discriminations to a system (or non-system) of no grades.

A fairly large proportion (37.5 per cent) of respondents favor adding pluses and minuses to the present grade discriminations. The general feeling is that such refinement would allow for greater equity. Two instructors suggested a numerical system. Two instructors favored a more limited application of pluses and minuses, one suggesting their use in basic courses, and another recommending their use in major courses only.

Some form of pass-fail system was favored by 40.6 per cent of those interviewed. However, many possible qualifications were mentioned. (The heterogeneity of the student body in ability and quality of performance seemed, for several instructors, to contraindicate a pass-fail system.) Of the qualifications mentioned, five faculty members would limit the pass-fail system to upperclassmen, five would limit it to elective, non-major courses, and four would endorse it only if there were smaller classes.

Three instructors interviewed suggested more radical changes in the system, two of them strongly urging written evaluations (with or without grades) as far more reliable indicators of performance.

Implications and Recommendations

Variability in Grading

Grades are presumed to have a universal, consistent, invariant meaning. Obviously, they do not. On a more modest level, grades should mean something that will generalize beyond a particular class at a particular time. Perhaps they do.

What is incontrovertible is that grade distributions show a vast amount of variability. This is true whether the comparisons are made between departments, between advanced and basic courses, between specific courses, or between sections of the same course. One class will produce a distribution skewed in one direction, another a distribution skewed in an opposite direction. For one kind of course, a failure rate, while varying somewhat, will remain relatively low with never more than a tenth of the class earning failing grades in any given semester.

Variability in grading as a kind of fine evaluation or judgment is expected. Our study, however, was focused on a gross measurement -- passing or failing grades -- with an implied judgment of adequate or inadequate performance. Variability in such a judgment demands further explanation.

What are some of the sources of this variability? First, pure chance must be acknowledged as one factor in the guise of the stroke of fate that places a group of exceptionally able students in one classroom, and a bevy of positive dullards in another. The extent of such chance variability can, however, be predicted on the basis of probability theory. In the absence of a known source of bias, other than possible selection through the rigors of each semester's course registration (where the more intelligent are able to get into the courses meeting at the most desirable times), one must presume chance to be a relatively minor source of variability.

The inherent difficulty of course material, the different levels of teacher effectiveness, the physical surroundings of the classrooms, international and national events, changes in selective service system, and many other factors also contribute in some small part to variability in grade distributions. Normal variability of subjective judgment is also a significant factor.

A particularly potent factor within a given college, group, or department is the presence or absence of a fair degree of explicit interaction between individual faculty members on the subject of grades. Studies have indicated that where grades are discussed, they tend to be less variable. This is true even when the discussion does not specifically set forth guidelines for grades, but instead covers institutional or course objectives and general standards of achievement.

The logical inference is clear. Much variability in grades is due to different concepts that instructors have of course objectives, of the function and meaning of grades, and of the institution's standards of achievement. An excellent illustration of the latter is the commonly recited account of the apocryphal instructor who cheerfully warns his students at the beginning of each term that the highest grade that they can possibly hope to receive is a "B." Most students, he tells them, will receive "C's"; "A's" are reserved for Nobel Prize winners, acclaimed scholars, and saints. In this instructor's class the mark or grade "B" therefore, means precisely the same thing as the mark or grade "A" would mean in most of his colleagues' classes.

There is a reasonable basis to infer, or at least suspect, that wide variability in grading may stem from mutually incompatible or even contradictory notions of objectives or standards. This inference is strongly supported by Juola's study of grading at Michigan State in which individual instructor's grades tend to correlate poorly with both mean GPA and scores of common (institution or course-wide) examinations. (Mean GPA in specific courses has a -.20 rank order correlation, with mean GPA secured by these students in other courses.)

Given the many possible sources of variability, an inescapable conclusion is that a major factor still remains individual instructors' differing concepts and perceptions, primarily of course objectives and of overall standards.

Basic vs. Advanced Courses

Our data clearly indicate that the grade distributions for basic courses differ markedly from those achieved in advanced courses. While this is not invariably true for all departments, it is certainly true for the majority of departments in our College. This difference can be ascribed to a number of major factors. One factor is the concept that the freshman year is an appropriate time for the weeding out of students who really should not be studying a given subject. This weeding out can either be consciously and deliberately done by the instructor, or it can be a process of "survival of the fittest" in a perfectly objective environment.

Another factor is the quite different perception that most instructors have of advanced courses. First, they are very much aware that students in these courses will be going on to graduate school where their admission will be conditioned on the level of their marks. Second, most instructors raise standards in these courses, and with the raising of the standards in terms of the kind of work that is to be done, they often become less rigorous in marking.

One should not, of course, ignore the factor of the student's becoming acclimatized to the college environment and gaining greater ability to complete his work successfully in a course. Nor should one overlook the self-selection process by which a student, even one who has done well in a basic course, who is not suited for a particular subject loses his interest and does not continue in the field. What this implies, of course, is that students in advanced courses may have a higher level of interest and motivation.

Although most of the instructors who were respondents in our study could explain the discrepancies between basic and advanced level courses, they were also somewhat surprised by the extent of this discrepancy. Actually, viewing the college experience as consisting of two distinct phases does, in a sense, represent a revolutionary view of what transpires in college.

Dissatisfaction with Grading

Few faculty people are satisfied with the grading system that currently prevails. What may well inhibit a change (although it is difficult to understand how this has also inhibited research in this area), is that the dissatisfaction is in two directions. Some faculty members are dissatisfied because grades are not sufficiently precise. They would like pluses and minuses to be added, a numerical system to be introduced, or some other method used to make grades much more precise. A somewhat larger group holds the contrary opinion that letter grades as now used are too precise. They suggest a number of systems, usually built around the pass-fail system, sometimes with a third grade included to cover the students who pass with a special distinction.

It would well appear, however, that the principal outlet for instructor dissatisfaction is occasional arguments between friends that sometimes erupts into a departmental meeting where some temporary agreement on course objectives and grading standards is achieved.

Structured Content Courses

A somewhat tentative inference that begins to emerge from our data (primarily the interviews) is that college courses are essentially made up of two quite distinct types or modes. One is the structured content course, such as mathematics or a foreign language, where a significant course objective is the student's acquisition of specific units of structured content. In these courses there are usually frequent examinations consisting of short answer questions in which the student is called upon to recall specific kinds of information.

Other courses are much less structured in their content. The acquisition of specific items of information is not nearly so important as new ways of thinking about and understanding the course subject. English composition, philosophy, psychology, and other social sciences and humanities, are of this nature.

Admittedly, the dividing line between the two modes is not distinct; there is much overlapping. What is even more confusing is that two instructors teaching the same course might well view the course from different perspectives. One might believe that the basic history course is primarily a structured content course, while another, less concerned with acquisition of facts, will view this as a less structured course aiming primarily at developing a way of thinking. If the distinction is valid, though, there is an interesting implication with regard to course grading. One might recommend for courses that are highly structured in content that more frequent examinations be given and that course grading be more precise. Here the arguments of those who propose adding pluses and minuses or moving toward a numerical system seem most justified.

In the less structured courses, where frequent detailed examinations of this kind are not possible, and where evaluation has to be more subjective, the possibility for obfuscating variability also increases. For these courses, it would seem that a pass-fail system, perhaps including a pass with distinction grade, would make a more valid evaluation.

While this solution, of course, raises almost as many problems as it solves, one could very well conceive of an effectively functioning system in which courses were identified by their instructors as either being structured content or less structured content. The grade system would depend on the designation of the course and the student would, through the course of his college career, essentially earn two different types of cumulative average grades.

The Need for Interaction

Without any changes in the current system, one specific recommendation, however, can be made. It is that departments and schools provide greater opportunity for the faculty to meet and discuss basic educational processes. The great variability that exists in grades indicate that, indeed, there may well be varying, even conflicting, concepts of course objectives, standards of achievement, and the overall college objectives. Again, where faculty meet to discuss some of these more abstract issues, there is specific evidence that grade distributions begin to become more consistent. It should be emphasized that no one who seriously studies the question of the meaning of grades proposes that grade distributions become identical. The plea has been that grades be sufficiently consistent so that the same letter grade, or same number, can be reasonably presumed to have a similar meaning. In a society and in an educational system in which grades, cumulative grades, and grade point averages are used to make crucial decisions about the future lives and careers of young men and women, this does not seem to be too much to ask. Ideally, this common meaning can best come about from the opportunity for the faculty to discuss openly and frequently the underlying questions. Dictation of any kind, certainly including dictation of an ideal or desired form of grade distribution, would be harmful to the educational process.

STUDIES OF COOPERATION IN THE CLAREMONT COLLEGES: ACADEMIC COOPERATION

Clifford T. Stewart
Director, Office of Institutional Research
The Claremont Colleges

In 1925, The Claremont Colleges began an experiment which was unique in American higher education. This experiment was the establishment of a cluster college or group of independent but cooperating colleges, which would provide superior intellectual resources for increasing numbers of students while maintaining the personal relationships of the small college. The goal was to develop in Claremont a group of small colleges somewhat of the Oxford type, which would be built around a library and which would have other central facilities or joint services. The cluster began by establishing the Claremont Graduate School and University Center which was charged with 1) providing graduate instruction, 2) developing new colleges, and 3) maintaining the common facilities for the group.

Today there are six colleges: Pomona College, Scripps College, Claremont Men's College, Harvey Mudd College, Pitzer College, and Claremont Graduate School and University Center. The six are independent institutions--each with its own president, board of trustees, students, faculty, and academic programs. The physical relationship of the colleges is shown in Figure 1.

Table 1 reveals eighteen activities in which all of the Claremont Colleges cooperate and another twelve in which combinations of two, three, four, or five cooperate. Cooperative activities with colleges outside of Claremont are not shown. I shall discuss one aspect of cooperation--perhaps the most crucial from an educational perspective--the joint planning of courses and programs among the colleges and the resultant cross registration or registration in a college other than the student's own college.

The purpose of permitting course registrations in other colleges is to make available to the student as wide a range of courses as possible. It also has certain economic implications for the participating colleges.

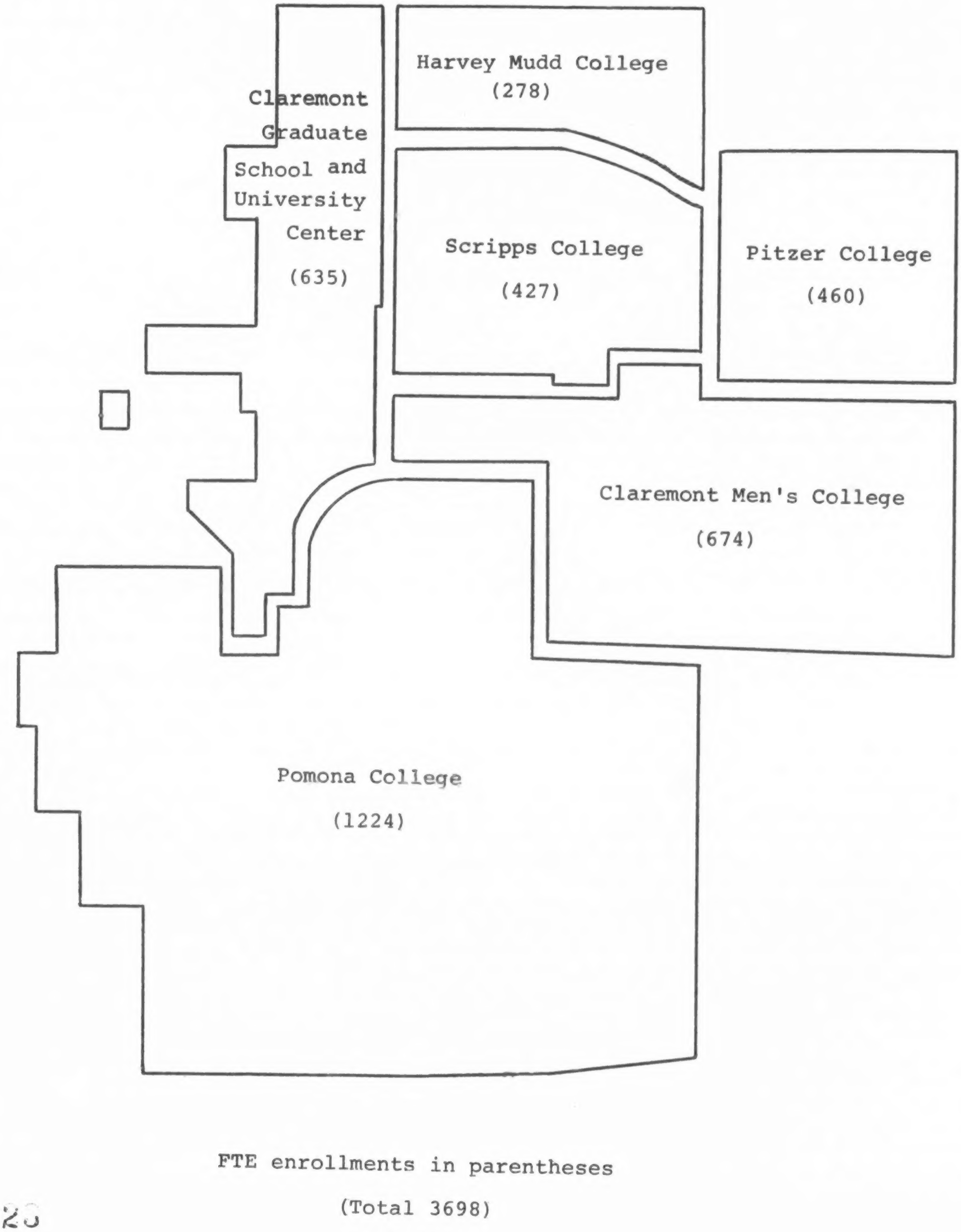
Some of the colleges have specific curricular emphases, whereas others are broadly based liberal arts colleges. For example, Pitzer College emphasizes the social and behavioral sciences; Scripps College emphasizes humanities; Claremont Men's College began with an emphasis in political economy (although all are liberal arts colleges.) However, not all of the colleges attempt to maintain program depth in all fields.

There are three types of courses which are involved in the cooperative activities of the colleges. First, there are the regular courses offered by each college, and most of these are open to students from all of the other colleges.

Second, there are joint programs and courses that are designed to involve students from several colleges. For example, the colleges offer a program in classics, but no one college wishes to bear the full expense of a university-size classics department. The four colleges with classics departments, therefore, cooperate to offer a full program in classics to students in all the colleges.

As another example, three of the colleges with a combined enrollment of 1600 have formed a joint science department. None of the three colleges emphasizes the natural sciences and therefore they have combined their resources to provide science offerings for their students. Nine faculty members on joint appointment with the three colleges make up the science faculty. Joint facilities are used,

THE CLAREMONT COLLEGES


Harvey Mudd College
(278)
Claremont
Graduate
School and
University
Center
(635)
Scripps College
(427)
Pitzer College
(460)
Claremont Men's College
(674)
Pomona College
(1224)

FTE enrollments in parentheses

(Total 3698)

FIGURE 1

TABLE I

Cooperative Ventures and Joint Facilities of the Claremont Colleges*

		CGS	CMC	HMC	PIT	POM	SCR
1.	Honnold Library	**1952	1952	1957	1963	1952	1952
2.	Business Office	26	46	57	63	26	27
3.	McAlister Center for Religious Activities and the College Church	59	59	59	63	59	59
4.	Service Shops	26	46	57	63	26	27
5.	Bridges Auditorium	30	46	57	63	30	30
6.	Telephone Service	29	46	57	63	29	29
7.	Health Service and Infirmary	66	46	57	63	27	27
8.	Psychological Clinic and Counseling Center	61	49	60	63	49	49
9.	Campus Security	46	46	57	63	46	46
10.	Garrison Theater	63	63	63	63	63	63
11.	Faculty House	55	55	57	63	55	55
12.	Office of the Provost	61	61	61	63	61	61
13.	Public Information Office	48	48	55	63	61	48
14.	Print Shop and Addressograph	62	62	62	63	62	62
15.	Challenge Campaign	65	65	65	65	65	65
16.	Cross Registration of Courses	26	46	57	63	26	27
17.	Faculty Exchange	26	46	57	63	26	27
18.	Office of Institutional Research	65	65	65	65	65	65
19.	Student Newspaper		57	57	63	63	57
20.	Heating Plant	26	48			26	27
21.	ROTC		46			46	
22.	Admissions		57	57			
23.	Joint Science Program		64		64		65
24.	Anthropology and Sociology				65	65	
25.	Classics		48		63	27	27
26.	Drama and Music		66	66	66		66
27.	Physical Education		57	57			
28.	Career Counseling				65		65
29.	Language Lab		60		64		60
30.	Forensics		66		66		66

*Does not include cooperative efforts with colleges outside of Claremont.

**The year listed is the one in which that college first took part in the joint effort.

and a new joint science building is being constructed to augment the existing facilities. Each of the two undergraduate colleges not involved in this program offers a full range of courses in the natural sciences. The joint science program has been formalized in an agreement which specifies the role of each college in appointment of faculty, financial responsibility, etc.

Included in this category of joint programs and courses are less formal arrangements for selected courses in mathematics or sociology, as examples. These courses are offered jointly, even though there are no overall joint programs in these areas. The faculties of the various colleges in certain disciplines meet and decide on courses to be offered jointly. Courses chosen are those which should be offered in the program but which one college would have difficulty filling each semester. Such courses may be at elementary, intermediate, or advanced levels.

A third category of courses consists of those taught by graduate faculty for undergraduate students. These are called intercollegiate courses and are taught in a variety of subject areas.

Regular courses enroll 1200 to 1300 students from other campuses each semester. Joint programs and courses enroll 100 to 200 students per semester and intercollegiate courses, another 100 to 300, depending on the number of courses offered in a given semester.

Approximately 10 per cent of all course registrations are in the types of courses mentioned above. Each semester one-third of the students take one or more courses off campus.

How does this academic cooperation work? What does it accomplish? It accomplishes its main goal of providing a wide range of course offerings and sufficient depth in the various fields to students in all the colleges. Students may take these courses as part of their major or may use them as elective subjects. It also gives the students the opportunity to choose professors from other colleges and to be in classes with students from other colleges. A student registers on his own campus for all the courses he wishes to take in any of the Claremont Colleges.

Are there any problems? There are many. Curriculum planning can become quite a problem since it may involve two to six colleges. Each college is free to develop its own curriculum and this, in fact, is what happens. However, there is usually much coordination among the deans and the faculties of the colleges in their continuing efforts to use available and new resources in the most efficient way. When a new course is to be added in any college or when a new faculty member is to be hired, this information becomes available to all of the academic deans and to the faculty members in the appropriate areas. This usually, but not always, eliminates duplicate courses being offered, and the hiring by several colleges of faculty members with the same specialties within disciplines. There are problems of coordination among the faculties of the various colleges concerning who will teach what next semester, who will be on sabbatical next semester or next year, and who will replace them.

All of this requires much time, but it provides an automatic and constant review of the curricula of the various colleges. The deans and the faculty are also constantly aware of the make-up of faculty in the various colleges and are therefore in a better position to make decisions concerning courses and staffing.

In certain laboratory, art instruction, and language courses there are a limited number of spaces or facilities, which determine the upper limit of class size. Therefore, registration must be limited in such courses, and this usually means the students from other colleges are denied registration in these courses. Fortunately these limitations do not affect more than a few students each semester. If the problem becomes more serious, then a different system of registration and/or a different distribution of course offerings may be required.

As in any college or university, courses occasionally become too large. This is potentially more of a problem at the Claremont Colleges since students register for all their courses (including those offered by other colleges) on their home campus. Therefore, class sizes are not fully known until some time after registration on all campuses has been completed. Since there are generally no limitations placed on class size, it sometimes becomes necessary to teach large sections. These sections are divided into smaller ones when faculty are available to teach them.

TABLE II

Cross Registration Data for Regular Courses and Joint Programs
Spring 1965-66

Registered at	Students of						
	CGS	CMC	HMC	PIT	POM	SCR	TOTAL
Claremont Graduate School	XX	22	3	0	9	6	40
Claremont Men's College	3	XX	41	126	38	85	293
Harvey Mudd College	2	10	XX	13	11	7	43
Pitzer College	3	65	17	XX	51	36	172
Pomona College	39	55	57	116	XX	155	422
Scripps College	9	91	24	61	68	XX	253
Total	56	243	142	316	177	289	1223
Balance for Regular Courses	-16	+50	-99	-144	+245	-36	
Balance for Joint Programs	+3	-39	-6	+50	-11	+3	
Balance for Regular Courses plus Joint Programs	-13	+11	-105	-94	+234	-33	

Table II shows the pattern of cross registrations in regular courses among the six colleges. One notices that Harvey Mudd, the smallest college and the one with the sharpest curricular emphases, and Pitzer, the newest college, send far more students to other colleges than they teach in return. Three hundred sixteen Pitzer students took courses in other colleges whereas only 172 students from other colleges enrolled in the regular Pitzer courses leaving a "deficit" of 144. However, Pitzer taught more students in joint programs and courses than it sent out, but still has a "deficit" of 94. Pomona College, on the other hand, has a positive balance.

Table III shows Pitzer College with less of a "deficit" for the regular courses and a positive balance when the joint programs and courses are added. Pitzer students take advantage of the wide variety of course offerings at the other colleges, the largest numbers of their students attending Claremont Men's College and Pomona College, the two largest colleges. Enrollments from other colleges at Pitzer are predominately in the joint science program, in anthropology, and in music courses. This illustrates that a new college can allow its students to register in other colleges for courses it does not offer, simultaneously attracting students from other colleges into strong programs in the social and behavioral sciences, the natural sciences, and music. This attraction is possible because Pitzer has concentrated to some extent on developing programs not emphasized by the other colleges. For example, only one of the other colleges offers work in anthropology. A strong program in this area is now available because Pitzer has planned this program with Pomona College and has committed its resources to securing excellent faculty in anthropology, sociology, and psychology. And this new, small college can attract good faculty because there are opportunities for them to teach students in these fields at several undergraduate colleges and, in some fields, at the Graduate School.

TABLE III

Cross Registration Data for Regular Courses and Joint Programs
Spring 1966-67

Registered at	Students of						
	CGS	CMC	HMC	PIT	POM	SCR	TOTAL
Claremont Graduate School	XX	14	2	4	16	13	49
Claremont Men's College	0	XX	39	97	34	72	242
Harvey Mudd College	3	10	XX	13	7	10	43
Pitzer College	5	49	28	XX	103	70	255
Pomona College	31	71	48	126	XX	182	458
Scripps College	9	84	19	68	93	XX	273
Total	48	228	136	308	253	347	1320
Balance for Regular Courses	+1	+14	-93	-53	+205	-74	
Balance for Joint Programs	-3	-3	-5	+63	-63	+11	
Balance for Regular Courses plus Joint Programs	-2	+11	-98	+10	+142	-63	

The students, of course, are highly in favor of cross registration, not only because it allows the opportunity to take courses not available on their own campus, but also because it provides the opportunity for interaction with students and faculty of the other colleges. Freshmen and sophomores are urged to take all of their courses at their college of residence, but they are not denied the opportunity to register in other colleges. Indeed, one-third of all off-campus registrations are by freshmen and sophomores.

Our Office of Institutional Research recently conducted a study jointly with the editors of the student newspapers of the five undergraduate colleges to determine student reaction to various aspects of the cooperative activities among the colleges. Student leaders on each campus were interviewed and questionnaires were sent to a sample of students on each campus. In addition to the previously listed advantages, some students reported taking courses at other colleges because of conflicts of courses on their home campus or simply because the time was more convenient. More comments were made concerning cross registration than about any other cooperative venture of the colleges. Students strongly support this type of academic venture; in fact, they want even more opportunity for cross registration. The policies governing cross registration are relatively liberal, but the students want them liberalized further. The deans of the various colleges have been aware of the students' wishes and are currently considering a revision of the policies in response to student requests.

The imbalance in some colleges from sending students to other campuses is not paid, but there is an effort made for the next semester or year to achieve a better balance between colleges. It generally has been true, for example, that Pomona College has been carrying a heavier load than it has been imposing on other colleges. However, by reference to Tables II and III one sees less imbalance this year than last. There are more students (more accurately, course registrations) at Pomona this semester than spring semester last year, but there are far more Pomona students enrolled in other colleges than was the case one year ago. In addition to this, Pomona students have enrolled heavily in the joint programs and courses supported by colleges other than Pomona. There is never an attempt to achieve an exact balance between any two colleges, but only an attempt to see that no one or two colleges have to support an excessive number of extra students.

In conclusion, the advantages that I have discussed in relation to cross registration have been academic advantages. The values accrue to the student, to the faculties, and to the college community. The current problems and foreseeable future problems in this area are largely mechanical. Other studies are underway which consider the financial aspects of the various cooperative efforts of the colleges. Our hope is that these studies will provide additional information which will be helpful in academic planning.

THE ROLE OF INSTITUTIONAL RESEARCH IN THE IMPROVEMENT OF INSTRUCTION

Fred Glassburner
Director, Institutional Research
Wisconsin State University--Platteville

Research is an essential ingredient of any systematic effort to improve instruction, and Institutional Research has a significant potential contribution to make in this effort, playing three essential roles in the instructional process. First, the improvement of instruction is a complex problem, calling for the systematic problem-solving procedures which characterize good research. There is need for the trained researcher's skill in identifying problems and in formulating and testing hypotheses. Second, improvement means change in what the teacher does; that is, it is change in the teacher's behavior. Change in behavior, by psychological definition, is evidence that learning has occurred, and one of the best methods of learning is through research. In fact, if one examines closely an analysis of the learning process, one will discover that it is remarkably like the outline for a research project. And thirdly, research lends respectability to educational experimentation and innovation. A professor might be reluctant to try alone that which he would willingly undertake as part of a research project. The word research itself has good connotations. "Trying it out on the dog" sounds much better when you call it research. Because of these views, the Office of Institutional Research at Wisconsin State University--Platteville has volunteered its services to the Improvement of Instruction Board, and has suggested a plan.

The Improvement of Instruction Board, made up of faculty and students, was created at the beginning of the current academic year. We took our offer of assistance and our plan to the Board's first meeting, last September. Newly created, the Board was faced with the necessity of choosing a course to pursue in undertaking its responsibilities. It gave our plan a cordial hearing, followed by an endorsement and recommendation that we proceed with it, with assurance of cooperation.

The Plan

The proposed plan uses the familiar problem-solving pattern: identify the problem, formulate hypotheses for solving the problem, and finally, test the hypotheses.

Identifying the problem was to be difficult. We were faced with difficulty at the outset, of course, to define improvement of instruction. There seems to be no standard definition of good teaching. Stanford University is doing research to try to define it, and they report that they expect to take about twenty years to do it. Nathan Gage, who is directing that research, says that one of the first obstacles they encountered was to define teaching. Not good teaching, just teaching. They finally settled on the following: teaching is what the teacher does in the classroom.

Because of the lack of agreement on what good teaching is, it logically follows that what constitutes improvement of instruction is a matter of opinion. The question is, whose opinion?

We decided that the most valid opinion, for our purposes, would be a consensus of the two groups most directly concerned: namely, the students and the faculty. Whether or not any such consensus really existed was uncertain, but we decided we had to try for it and see what would develop.

The first data-gathering device was an open-ended questionnaire, which said, simply: "The improvement of Instruction Board is seeking suggestions for the improvement of instruction. Please write below any suggestions you may have." We sent it to a sample of 150 faculty and students in approximately equal proportions.

At the time of writing, the responses are coming in, and are being edited and converted to simple action statements appropriate to a rating scale. Here are some examples of suggestions received thus far:

"Get teachers more interested in what they're teaching." (From a student.)

"Abolish compulsory class attendance." (From a faculty member.)

"Encourage more student-faculty interaction." (From a student.)

"Start a committee to develop an overall philosophy of education and plan for the growth of academic functions." (Faculty.)

"Reduce class-cutting by instructors." (Student.)

I have picked those examples more or less at random, and not because I consider them particularly outstanding. How they will fare in the ratings is anybody's guess.

After receiving all suggestions, we shall probably have a list of fifty items, which will be set up to be rated as follows: 3 equals a very good idea; 2, merely a good idea; 1, a fair idea; and 0, the idea has little or no merit.

We plan to submit the rating scales to random samples of fifty faculty and fifty students and shall strive for one hundred percent response. To the faculty, we shall send a form to return to us separately after completing and returning the instrument, thus preserving the responder's anonymity--but enabling us to know who has not responded, so we may follow with further urging.

As for the students, we plan to assemble the subjects in a classroom and administer the rating scales, since previous experience with mail-in returns has not been too good. We shall also use the opportunity to administer two other instruments in connection with other research.

Along with the rating scales, we will submit a list of statements of educational philosophy, to which we will ask them to indicate agreement, disagreement, or no opinion. We are deriving the statements from the suggestions for improving instruction, by what we hope is an inductive process. For example, from the first suggestion quoted above ("Get teachers more interested in what they're teaching"), we might induce the following: The active interest and enthusiasm which a teacher displays tells more about how he really feels about his subject than what he says about it, and the student's own developing attitudes toward the subject will be influenced accordingly."

Hence, we are trying to develop a philosophy of education by consensus, basing it on statements which have been supported by a significant majority of faculty and students. Therefore we shall have a stated philosophy which can become a living document to guide us in evaluating existing practices and designing future innovations. Too often institutional philosophies are made up of cliches to which we give lip service, but which are not reflected in practice.

When the responses to the rating scales and the philosophy statements have been completed and tabulated, we hope that the resulting ratings and percentages will indicate promising areas and directions for innovation. If they do, we can then solicit volunteers from faculty and students to participate in planning and carrying out curricular experiments, carefully designed with rigorous controls and measuring instruments to ascertain whether or not the tested ideas actually achieve the desired ends. In short, this area is where we test hypotheses. The innovations undertaken will be regarded as tentative, pending the outcome of the experiments. This procedure is in contrast with the common practice of following educational fads and initiating innovations based on attractive but questionable premises, which become institutionalized and permanent without a rigorous evaluation of their merits ever having been made.

It may be that using a research approach to curricular innovation, with all that approach implies in the way of the tentative nature of hypothesis and the withholding of final judgment until all evidence received, will give the teacher greater freedom of action to break the crystallized tradition. Negative reaction to a proposed new idea is usually in the form of the assertion that it will not work, for a number of presumed reasons. The researcher can respond that perhaps it will not, but we will never know until we try. If it does not, we can drop it. This is a much more tenable position than arguing that it will work, for a number of other reasons, also presumed. Also, if we are successful in having an idea adopted merely on the basis of our faith, our enthusiasm, and our powers of persuasion, we then find ourselves in the dilemma of either maintaining the fiction that the idea worked if it really did not or admitting to a large error in judgment. The former choice is taken more often than the latter, with the result that we are ritualistically performing many educational practices which are hallowed by time and usage, perpetuated by ego-involvement and vested interest, but which never really worked as intended. A research approach to the improvement of instruction will tend to prevent such dilemmas and invalid rituals.

One important outcome we anticipate is the effect of the research process upon the participating faculty member, who presumably will take part in planning the experiment and in carrying it out. Researchers know that every investigation is a learning experience for all involved in it. We can assume that the same will be true for faculty members. Moreover, in the case of disciplines other than teacher-training, our research will be in the teaching-learning process. The disdain with which professors in other fields often view the discipline of education has, I think, been fairly well documented. It seems somewhat paradoxical that proud men should look upon their own craft, which is teaching, as not being worthy of study and investigation.

In any event, we hope to bootleg some educational research into the lives and consciousness of people who have not thought much about it heretofore, and perhaps in the future they will view the teaching-learning process a little more critically and analytically.

The part to be played by students poses some special problems. If they are aware that they are participating in an experiment, Hawthorne effect seems inevitable, especially in the case of radical departures where the use of volunteers seems called for. We have no intention of hiding our light under a bushel, and plan to foster general interest by means of as much favorable publicity as we can decently employ. Students will therefore be well aware that they are in an experiment, which means that if the measured effects of the treatment prove favorable, we will not know whether it is really the treatment, or just Hawthorne effect. We will have to admit that although the patient improved the operation was a failure. If this failure occurs, perhaps we shall extend the period of the experiment another year or two, and continue to measure and detect any decline in the apparent benefits. If worst comes to worst, and Hawthorne effect is all we can receive, we must turn our efforts to concocting additional experiments.

Finally, there is the anticipated benefit of improvement in faculty-student relationships. On the national scene there seems to be growing disaffection between students and administration. The Association for Higher Education has launched a study of student participation in campus governance as a means of counteracting that trend. Perhaps by enlisting students as active participants both as planners and subjects, we shall take a step toward casting them in the role of partners in the educational enterprise rather than antagonists.

Summary

To recapitulate, I have described a plan for an institutional research project for the improvement of instruction, which is now in progress at Wisconsin State University--Platteville. It began with an open-ended questionnaire soliciting suggestions from faculty and students, from which rating scales are being constructed to measure the support for each suggestion. Also, statements of educational philosophy were distributed for which students and faculty will be polled to ascertain the extent of agreement or disagreement. On the basis of the ratings and philosophy thus derived, instructional experiments will be designed, with appropriate controls and measurements of outcomes. It is expected that participation in the project by faculty members will give them a better understanding of the teaching-learning process.

RESEARCH IN THE INSTRUCTIONAL PROCESS: A MODEL AND A PROPOSAL

Peter P. Grande
Associate Professor of Education
University of Notre Dame

The present study is an exploration of the relationship between specified learning outcomes and clearly defined classroom conditions in a "live" classroom learning situation. Attention has been accorded both to the need for information contributing to a "relevant descriptive base for the study of teaching" and to the need to explore "more tightly defined learning situations." The construct "classroom conditions" is used in this study to denote certain interpenetrating factors which "mix," optimally or otherwise, to produce a distinctive classroom climate. "Learning outcomes" is used to denote certain effects of classroom learning conditions which are potentially important in exploring specific modes of "reacting" to encountered situations. "Student characteristics" is used to denote certain elements potentially contributory to the efficacy of the classroom learning process: personality, intelligence, customary school achievement level, socioeconomic status, and family interaction pattern. Finally, the construct "'live' classroom situation" is intended to convey that this study was executed not in an experimental laboratory analogue or approximation to a "real-life" classroom, but rather within such a classroom itself, with all the contaminating and confounding effects which "real-life" situations introduce into educational research.

The variables which form the foci of interest in this investigation, as important parameters of classroom learning conditions and as potentially significant influences on learning outcomes, are: (1) instructional technique, (2) instructional group size, and (3) an ipsative variable, conceptualized as "teacher style," or the specific fashion in which instructional technique is mediated through the characteristic behavior pattern of the individual teacher. Learning outcomes which form variables of focal interest in assessing ways of reacting to encountered classroom situations are: (4) student mastery of subject content, (5) resultant level of student mental processes, (6) student attitude toward and satisfaction with the specific "mix" of classroom conditions, and (7) student participation in the classroom learning situation. Student characteristics which form variables of interest in determining contributory influences upon the specific "mix" of the classroom learning situation include: (8) personality factors, (9) intelligence, (10) customary school achievement level, and (11) socioeconomic status. Hence, the principal research questions which this study seeks to explore are:

I. How do these specified classroom learning conditions, operating alone and in interaction with each other, affect specified learning outcomes?

II. How are these specified learning outcomes related to each other?

III. How are these specified learning outcomes related to specified student characteristics?

Conceptual Foundations

According to Wallen and Travers (1963, Gage, pp. 448-505), patterns of teacher behavior in the classroom, or instructional modes, originate from a variety of sources, the least productive of which have been those empirically derived. Wallen and Travers suggest that the teacher's classroom behavior is derived from (1) teaching traditions, i.e., a teacher teaches as he was taught; (2) philosophic traditions, i.e., a teacher behaves in accordance with the doctrines of, for example, Freobel, Rousseau, Dewey, or Montessori; (3) social learnings in the teacher's background, i.e., a teacher behaves to reinforce pupil behavior consonant with, for example, middle-class or democratic values; (4) the teacher's own needs, e.g., a teacher adopts and adheres to the lecture method because he needs an outlet for his self-assertive tendencies; (5) conditions within school and community, e.g., a teacher conducts his classes to produce the formal and highly disciplined behavior which corresponds to the pattern favored by administrative officers; and (6) models for teacher behavior emanating from scientific research.

Specifying desired learning outcomes is no facile task for the educator, in the face of a number of cogent but competing rationales. One rationale focuses on the developmental needs of the child as providing the soundest basis for selection of learning outcomes (Jersild and Fehlmann, 1943), while others focus on the psychological needs of the student (Kotinsky and Coleman, 1955), on consistency with social realities (Joyce, 1961), or on the facilitation of intelligent self-direction by students (Smith, Stanley, and Shores, 1957).

Intersecting such rationales for specifying desirable outcomes in classroom learning are the problems of establishing and operationalizing suitable criterion measures. As Taba (1962, p. 266) maintains: "A clearer distinction between the content of the curriculum and the learning experience or the processes which students employ in dealing with content would be helpful in classifying problems of selection as determining which criteria to apply to which aspect of the curriculum." This point of view suggests a need for research aimed at a clear delineation between curriculum content and its impact, mediated through specified classroom learning conditions, upon student mental processes, as one potentially significant learning outcome.

Gage (1964, pp. 268-286) urges that "theories of learning will have greater usefulness to education when they are transformed into theories of teaching," for "while theories of learning deal with the ways in which an organism learns, theories of teaching deal with the ways a person influences an organism to learn." In this approach, the emphasis shifts from a consideration of the characteristics of the learner and of the learning process to a consideration of the characteristics of the teacher and of the teaching process. In the design of the present study, Gage's view that "changes in education must depend upon what the teacher does" is accorded paramount importance, although the design of this study remains cognizant that what the teacher is and does intersects in the learning climate with what the learner is and does in influencing learning outcomes.

Wallen and Travers (Gage, 1963, p. 464), who uses the terms "teaching methods" and "patterns of teacher behavior" interchangeably, note that "most prescribed teaching patterns have been influenced much more by philosophical traditions, the needs of teachers and of professors of education" than by empirically-based rationales. Such an assertion implicitly questions the conceptual and operational efficiency of teaching methods which have sought to compare the effectiveness of one "method" with another. It serves also to underscore Ginther's (1964) admonition that more useful research on teaching-learning processes awaits "more tightly defined learning situations." Wallen and Travers (Gage, 1963, p. 493-494) appear to concur when they declare that "research on teaching methods which will contribute to an organized body of scientific information requires that teaching methods themselves be designed systematically" and that "research workers must surely go back, take stock of their position, and realize that the starting place must be the systematic design of teaching methods."

An effort has been made to systematically design the two experimental teaching methods employed in the present study in consonance with Ginther's (1964) model for instructional analysis. This three-dimensional schematization is founded upon a programmatic approach to instruction. Its first dimension, programming, represents a bi-polar conceptualization of the manner of control over classroom verbal interaction exercised by teacher behavior, specifically employing the polar constructs "errorless" vs. "dialectical." The "errorless" pole is characterized by a set of conditions for the control of instructional verbal interaction in which students are led through a series of programmatic steps so devised as to retard, or at least severely inhibit, "mistakes" on the part of the learner. Accordingly, the learner is permitted little or no freedom in his classroom behavior, since this polarity is implemented through a highly structured question-and-answer approach. The "dialectical" pole, on the other hand, is characterized by a set of conditions in which students are permitted considerable freedom in the classroom learning situation to stimulate them to come to an understanding of the matter to be learned through "discovery" and to correct deficiencies in information, misunderstandings, or inappropriate use of information. Considerable freedom is also accorded the learner to formulate and refine original ideas. This bi-polar programming dimension appears amenable to operationalization through instructional materials, as well as through teacher behavior pattern or instructional technique, although the present study operationalizes these polarities primarily through the modality of instructional technique.

Errorless and dialectical methods are illustrated in the instructional materials used in the present study, which consist of teaching units in English composition so highly detailed as to control teacher classroom verbal behavior. These units were developed at the University of Chicago's Center for the Cooperative Study of Instruction (Rippey, 1966). In classes instructed by an "errorless" technique, students are provided information in small incremental steps and are questioned immediately about this information to reinforce learning. An optimal errorless situation occurs when all students come to a consensus concerning a correct response, which is, incidentally, provided in each interaction by the teacher. In classes instructed by a dialectical technique, students are stimulated to ask and answer questions and even to assume control over the direction of class discussion. In the dialectical technique, student responses cannot be anticipated with the certitude available in errorless technique. Dialectical classroom discussions may well stray from the focal topic or even from the subject content. In this event, the teacher guides the discussion back to the matter at hand.

In addition to "errorless" and "dialectical," a third instructional "technique," termed "teacher-uncontrolled," is investigated in this study. As the term implies, the teacher-uncontrolled situation contains neither internal nor external restrictions over teacher behavior; in this sense, the teacher-uncontrolled modality represents the teacher's typical behavior pattern uninfluenced by experimental treatment methods, and is thus analogous to what is termed "traditional" teaching methods in many educational research investigations.

While the manner of control of verbal interaction in the classroom may be expected to vary in errorless and dialectical techniques, experimental investigation is needed to determine whether these differential patterns of verbal behavior, especially within the context of varying class size, are associated with differential learning outcomes.

Although the issue of instructional group size has been, and continues to be, a matter of both professional and public concern, more than six decades of educational research have failed to provide definitive information about what changes, if any, obtain in teaching-learning processes and in learning outcomes when class size varies. Conclusions from research on instructional group size in a variety of subject-matter areas, in fact, have displayed an amazing uniformity in reporting that achievement in skills measured by standardized tests is not significantly affected by class size. In a pioneering effort, Rice (1903) studied language learning in elementary school, while Corman (1909), Backman (1911, 1913), Boyer (1914), and Elliott (1914) were interested in promotion rates, and Breed and McCarthy (1916) and Whitney (1929) in spelling. Stevenson (1922) reported that small classes proved superior only for slow students, but Averill and Mueller (1925) found that pupils in small classes (12 students) showed greater gains in reading achievement than those in larger classes, whether or not they were slow students. More recently, Otto (1954) investigated large classes, with a median size of 37, and "small" classes, with a median size of 23. He reported that, while the learning environment seemed more effective in "small" classes, there were no significant differences in subject matter mastery, or achievement, between small and large classes. Whether a class of 23 pupils can, in fact, be appropriately regarded as "small" is a moot point. However, Spitzer (1954), investigating reading comprehension, study skills, language skills, and arithmetic skills, found no differences between large and small classes when pupils' gains were measured by traditional achievement measures.

In contrast, extensive research conducted at the Institute of Administrative Research, Teachers College, Columbia, prompted McKenna (1957, p. 437) to report that creativity is accelerated in small size classes and that "promising new procedures are more rapidly adopted in small classes." Further, McKenna reported that teachers of small classes came to know more about individual pupils and that there was a greater tendency to individualize instruction. More group work, in a more informal atmosphere, characterized the small classes. Enrichment materials were typically used in small classes, while the dominant mode of information exchange in large classes was reliance upon the textbook.

Hence, one may provisionally conclude that class size contributes little to desirable learning outcomes when the criterion is subject mastery as measured by a standardized achievement test. However, when criteria other than standardized test-measured scholastic achievement are employed, instructional group size appears to emerge as a more potent variable in classroom learning. Thus, it becomes important to conceptualize learning outcomes to include criteria which are regarded as desirable educational results in addition to achievement as assessed through standardized tests. Further, it is necessary that operational definitions of class size in educational research approximate, or at least not vary significantly from, the common meanings attached to such concepts in current educational practice. In the present study, small classes are those whose size is 10 or fewer students; medium, those whose size is approximately 25; and large, those whose size is 40 or more. It is anticipated that these purely quantitative distinctions will be accompanied by corresponding qualitative distinctions in terms of distribution of participation and the nature of student-student and student-teacher interaction (Thomas and Fink, 1963), and the assessment of group process in the classroom learning situation is an important goal in this study.

Learning has been customarily defined as behavior change, following, and perhaps resultant from, experience. The effectiveness of the classroom teaching process is generally assessed through measurable changes in learner behavior, customarily a change in mastery of subject matter content, as a result of participation in, or exposure to, particular learning experiences. In this approach, the assessment of the effectiveness of the teaching-learning interaction remains indirect, inferred from evidence gathered about its presumed effects in pre-post achievement measures.

Another, more direct, avenue for the exploration of learning outcomes is suggested in the pioneering studies of student classroom experiences by Bloom and Broder (1950), Bloom (1954), and Schulz (1951). These investigators sought to examine the learning experiences of students by exploring directly precisely what students experience during the classroom transaction, usually by investigating student thought processes during given class interactions. Early classificatory systems for categorizing student experiences employed rather gross categories, such as "recall" vs. "higher mental processes" (Tyler, 1936). These categories were considerably refined and rendered more discriminatory in Bloom's (1956) taxonomy of educational objectives, which provides a system for the classification of learning outcomes in the cognitive domain.

Learning outcomes are approached from both vantage points in the present study, which assesses learning outcomes (1) indirectly through the pre-post treatment administration of achievement measures and (2) directly through the analysis of student mental processes. In addition, student satisfaction with, and degree of participation in, the learning situation under varying sets and "mixes" of classroom conditions are explored as potentially significant learning outcomes.

Research Design

This study is to be executed during the second semester of the 1966-67 school year in three secondary schools of the South Bend (Indiana) Community School Corporation. Subjects in this investigation are to be 935 students enrolled in a sophomore (tenth grade) English course stressing skills in composition. In every sense, this study is to be conducted during "live" classroom situations, with students of varying levels of ability and motivation who are actually enrolled in a course for credit offered by the secondary schools which they normally attend, and with all the contaminating effects attendant upon a "real-life" classroom situation. Instructors participating in this investigation are four properly certified English teachers employed by the public schools who, until the time of their participation in this study, had neither knowledge of, nor familiarity with, errorless and dialectical instructional techniques. Indeed, they are not to be familiarized with these techniques until after the first phase of the study, which focuses on teacher-uncontrolled instruction, has been completed. In this sense, the present investigation attempts also to determine the extent to which teacher-uncontrolled behavior can be modified and molded into newer, non-traditional patterns, with obvious implications for teacher education.

The 935 student-subjects in this study constitute the entire population of the sophomore classes in the three schools in question. Prior to, and after, the experimental phase of this study, they have been and will continue to be instructed in classes of from between 24 and 35 members, according to techniques which are either ipsative to individual teachers, or at least unspecified. For purposes of this investigation, students are to be randomly reassigned from their customary sophomore English class to experimental classes of varying sizes, within the restrictions imposed by scheduling and space considerations. Each participating teacher is to instruct in each of the three experimental classroom group sizes according to each of the three instructional techniques.

Prior to the initiation of the first experimental phase, measures will be obtained of student (1) personality factors, through Cattell's High School Personality Factor Questionnaire; (2) intelligence through the Otis Quick-Scoring Test of Mental Ability; (3) customary school achievement level, through conversion of each student's rank-in-class to a standard score in accordance with procedures suggested by Educational Testing Service (1958); and, (4) family socioeconomic status, through classification of the principal breadwinner's occupation on the Hamburger (1959) Occupational Rating Scale. Prior to, and at the completion of, each of the three experimental phases in this investigation, measures will be obtained of (5) student subject mastery, through Rippey's (1965) Criterion Reference Test in English; (6) student attitude toward English, through the Silance-Remmers (1961) Generalized Attitude Scale; and (7) satisfaction with the "mix" of the specific classroom conditions to which each student has been exposed, through an attitude scale now in the process of construction for use in this investigation.

At the conclusion of the first, fourth, seventh, and tenth class session during each of the three experimental phases in this study, measures will be obtained of (8) student mental processes, through the technique of stimulated recall (Bloom, 1954). And finally, following the conclusion of the experimental phases in this study, measures will be obtained of (9) each student's subjective evaluation and critique of the specific "mix" of classroom conditions to which he has been exposed in relation to the "mix" he has typically experienced during his secondary school career.

During each experimental phase, teacher behavior is to be evaluated by project staff observers trained in classroom interaction process analysis, through the use of the Amidon-Flanders (1963) technique. Observations are to be utilized to insure conformity of teacher behavior to errorless or dialectical behavior pattern, by providing feedback from observers to teachers at the close of each session during the second and third experimental phases. Additionally, observations will be utilized to compare behavior patterns during the teacher-uncontrolled experimental condition with those emitted during errorless and dialectical conditions.

Each experimental phase occupies two weeks. Within the first two-week period, all four participant teachers will instruct approximately one-third of the 935 student subjects according to the teacher-uncontrolled pattern. A one-week interval follows this first phase, coincidental with the spring school vacation period, during which participant teachers will be given instruction and supervised practice in errorless and dialectical instructional techniques by experts in each serving as project consultants. Within the second two-week experimental phase, two participant teachers will instruct approximately one-sixth of the student subjects according to errorless technique, while the other two participant teachers will instruct another sixth according to the dialectical technique. Within the final two-week period, the two participant teachers who had instructed according to errorless technique during the second phase will instruct one-sixth of the student subjects according to dialectical technique, while their opposite numbers will instruct one-sixth according to the errorless technique.

Statistical Treatment

Data collected in this investigation in consideration of the first research question, which inquires into the effect of specified learning conditions upon specified learning outcomes, will be analyzed through analysis of variance procedures for three bases of classification, in a 2x2x2 matrix. Bases of classification are (1) class size (small, medium, large); (2) instructional technique (uncontrolled, errorless dialectical); and (3) teacher style (ipsative characteristics of teachers A,B, C, and D). Hence, the sources of variance to be analyzed are:

1. Class size

2. Instructional technique

3. Teacher style

4. First order interaction between class size and instructional technique

5. First order interaction between class size and teacher style

6. First order interaction between instructional technique and teacher style

7. Second order interaction between instructional technique, class size, and teacher style.

Into analysis of variance tables in a 2x2x2 matrix will be arrayed measures of student subject mastery, student attitude toward English, student satisfaction with the learning experience, and student mental processes.

Data collected in consideration of the second research question, which inquires into the relationship between learning outcomes, will be treated according to correlation techniques.

Data collected in consideration of the third research question, which inquires into the relationship between student characteristics and specified learning outcomes, are to be treated according to a variety of correlational techniques for single, partial, and multiple correlation for continuous and dichotomous variables.

References

Corman, Oliver P., "Size of Classes and School Progress," Psychological Clinic, Vol. 3 (December 15, 1909), 206-12.

Gage, Norman L. (Ed.). Handbook of Research on Teaching. Chicago, 1963.

Gage, Norman L., "Theories of Teaching," Theories of Learning and Instruction, (Ed.), Ernest R. Hilgard, LXIII (Chicago, 1964), 268-86.

Ginther, John R., "Conceptual Model for Analyzing Instruction," a paper presented at the Rochester Conference on Programmed Instruction in Medical Education, June 1964 (Chicago, 1964).

Jersild, Arthur T. and Charlotte Fehlmann, "Child Development and the Curriculum: Some General Principles," Journal of Experimental Education, XII (1943). 130-42.

Joyce, B.R., "Science, The Curriculum and the Young Conscience," Elementary School Journal, LXI (1961), 211-16.

Kotinsky, R. and J.V. Coleman, "Mental Health as an Educational Goal," Teachers' College Record, Vol. LXI (1955).

Rice, J.M., "Educational Research: The Results of a Test in Language," The Forum, Vol. 35 (October, 1903), 269-93.

Rippey, R.M., Errorless and Dialectical Units in English Composition. Chicago, 1966.

Smith, B. Othanel, William O. Stanley, and J. Harlan Shores, Fundamentals of Curriculum Development. Younkers, New York, 1957.

Taba, Hilda, Curriculum Development--Theory and Practice. New York, 1962.

INSTITUTIONAL OBJECTIVES AND PATTERNS OF STUDENT CHANGE

Arthur W. Chickering
Project Director, Project on Student Development In Small Colleges
Plainfield, Vermont

Two interrelated trends are rapidly gaining momentum in higher education. First, the relevance and effectiveness of the nation's colleges and universities increasingly has been called into question by students, college faculties and administrations, and governmental agencies. The Berkeley riots, the National Conference on Student Stress, the Campus Environment Studies of the United States National Student Association (supported by the National Institute of Mental Health)(9), the Conference on Innovation in Higher Education and the follow up meetings supported by the Office of Education, (11) and the establishment of the Union for Research and Experimentation in Higher Education are only the white caps topping a ground swell of increasing magnitude.

Second, as the establishment of the Association for Institutional Research and its rapid growth in membership attests, research concerning higher education seems to increase geometrically each year. Further, more and more of this research is directly addressed to institutional impact, to trying to discover relationships between institutional policies and practices and student development.

These two conditions, along with other pressures, are leading colleges and universities to confront more forthrightly three major questions:

1. Do our students change in college and is that change congruent with our objectives?

2. If change occurs, when and where does it take place?

3. Are there developmental principles which apply to student change in college which could help guide decisions?

The studies reported here are relevant to these questions as are the procedures by which the institution came to grips with them. Like current explorations of another space, two stages were required to get into orbit: first, re-definition of objectives; second, examination of student change.

Process

The process began with re-definition at a more concrete level of two major institutional objectives, the development of independence and of purpose. First, there was general discussion in faculty meetings concerning the definition of independence and what behaviors represented such development. Then each faculty member named five students who best represented high levels of independence, as he conceived it, and described the criteria underlying his choice. A faculty committee received these documents and pooled the criteria submitted, generating a definition which was returned to the general faculty meeting for futher discussion and modification. The Coordinator of Evaluation compared students most frequently nominated with their non-nominated peers and found significant differences between the two groups on a battery of tests and inventories administered to all at the end of their sophomore year. These differences were congruent with the major dimensions of the definition.

A similar procedure was followed to clarify the meaning of development of purpose. Again, nominated students and their non-nominated peers were compared and again significant differences were found congruent with the definition.

This work accomplished two aims. Institutional objectives had been translated into more concrete terms, and particular measures which characterized more highly developed students had been identified. Thus, assessment of the development of independence or of purpose could be undertaken through test-retest studies of first, second, and fourth year data. Furthermore, it would be possible to see whether development occurred primarily during the first two years, the last two, or rather evenly over the four year period. (A report of these test-retest studies is also available in the context of a general report on this experiment (3).)

These definitional activities yielded six major sub-categories of behaviors and characteristics which were labeled: Goal Directedness; Personal Stability and Integration; Venturing; Resourcefulness and Organization; Full Involvement, Motivation and Persistence; Interdependence. But it seemed wise to push further, to see whether objectives could be further detailed, and to see whether the timing and patterns for such development could be more precisely described. To these ends the faculty studies of student records were undertaken.

For each of these six variables questions concerning specific behaviors or attitudes represented in the four year records of the 1964 graduates were posed. Because Goddard uses a system of written self-evaluations and instructor comments rather than grades and examinations, and because the non-resident work term, the on-campus program, community participation, and extracurricular activities are subject to evaluation as well as academic study, student records are rich in material suitable for this kind of analysis. These records were rated for each semester on each question. Teams of four or five faculty members each assumed responsibility for one or more of the six variables to be assessed.

Preliminary discussion and trial ratings of several students helped create common standards for a scale from zero to ten with five set as the general expectation, average, or norm. The rating tasks were then divided so that each student was rated independently by two faculty members on each question. A student's semester score on a given variable was the pooled ratings on questions relevant to that variable. Faculty members were asked to rate a single semester for several students before returning to rate another semester for a particular individual. It was also suggested that semester reports be drawn at random from a student's folder. It was hoped that these practices would minimize:

1. The halo effect from one semester to another for a particular person;

2. Systematic distortions arising from moving in an ordered fashion from beginning semesters to later semesters;

3. Systematic distortions arising from subtle changes in the rater's standards over time.

The ratings for each of the six variables which resulted from this process were then standardized and subjected to multiple discriminant analyses (4) to discover whether change had occurred, and if so, which variables carried greatest weight. The actual questions posed for each variable are given as the results are discussed.

Results

Statistically significant change (beyond .01 level) was reflected by the multiple discriminant analysis, and covariance analyses of the scales individually revealed significant change (.01 level) on each. Weightings on the discriminant analysis indicated that ratings on Goal Directedness changed most, followed by Personal Stability and Integration, Venturesomeness, Resourcefulness and Organization, Full Involvement, Motivation and Persistence, and finally, Interdependence. Examination of each reveals the different timing and patterns of change which occurred.

Goal Directedness

The faculty described the purposeful student as one who had conscious and fairly well defined goals meaningful to himself. He had developed an increasing ability to see the relationships between his purposes and other aspects of his life, and his work increased in focus and depth through its relationship to his goals. The questions addressed to the records for the rating were:

1. Does the student plan his program with reference to a clear goal or purpose?

2. Are courses or independent studies evaluated in terms of their helpfulness or contribution to a larger purpose?

3. Are objectives for study explicitly related to more general plans or purposes?

4. Is there recognition of gaps in knowledge or skills in relation to purpose?

5. Are efforts made or plans formulated to deal with gaps or weaknesses?

6. Are there general expressions of feeling lost, at loose ends, without any purpose or direction? (Reverse scoring)

7. Do plans for the Nonresident Work Term reflect concern for some general plan or purpose?

8. How solid does the final commitment seem?

Averaging of the ratings of these eight questions, from two independent raters, yielded a score for each semester for each student.

The increase on this variable is substantial and steady, with the point of sharpest increase coming with the fourth semester. Goddard has Junior and Senior Divisions, and one applies for, and is accepted to, the Senior Division on the basis of his prior work. The Senior Division application requires a description of the work to be undertaken during the last two years. This hurdle, necessitating further specification of what one is to do and thus clarification of plans for the future, is probably responsible for the increase reflected at this time.

Personal Stability and Integration

Next in order of magnitude of change came Personal Stability and Integration. Both the independent student and the purposeful one were described by the faculty as having a higher level of stability and integration than their peers. They knew the kind of person they wanted to be and had a sense of balance and perspective. They tended to see things whole and with a well ordered set of values. They had sorted out what was important to them and were aware of their own strengths and weaknesses. They were relatively at ease about problems concerning academic work, future vocation, marriage and family life. It is not so much that such problems do not exist for them, or that they necessarily have resolved the problems, but that their level of anxiety and concern about such things is relatively low, and their comfort and confidence in their present state with regard to such matters is relatively high. The questions in this case were:

1. What is the student's level of reliability and responsibility on work programs in relation to other responsibilities undertaken?

2. What level of personal stability and integration is reflected by student

comments in relation to self, or self-development?

3. How comfortable is the student about the kind of person he is, or about the kind of person he was during the semester? How comfortable about his own past behavior is he?

Once more there is a picture of substantial and fairly regular change. It is interesting that the first semester ratings for students on this variable are considerably lower than for any of the others. We realize that the first semester particularly presents a constellation of adjustment problems which are frequently quite upsetting, and these ratings are congruent with that knowledge. It is also interesting to note that in addition to a sharp rise from the first to the second semester, there is also another acceleration in the seventh semester. This may suggest that the transition from college to the adult world has already begun, and that with the imminence of the senior study and with graduation impending a step is taken toward a higher level of personal organization and integration. This pattern also supports Mervin Freedman's suggestion that "Perhaps we should think of a developmental phase of late adolescence, beginning at some point in high school or prep school and terminating around the end of the sophomore year in college; followed by a developmental phase of young adulthood that begins around the junior year and carries over to a yet undetermined extent in the alumni years." (6)

Venturing

Venturing was the label supplied for the student who is open to experience, willing to confront questions and problems, to discover new possibilities, to disagree and be autonomous, and to initiate things for himself. The questions used to make connections with the records were these:

1. How much does the student speak up in class? How ready is he to express his own ideas and join the battle? Does he brood and maintain a stoic silence or does he externalize his feelings and ideas?

2. How easily does he communicate with the instructor? How free is he to disagree?

3. To what extent does he engage in study or other activities to tackle perceived weaknesses or liabilities?

4. How frequently does he speak of lack of self-confidence, of fears which restrict his activities? (Reverse scoring)

5. To what extent does he seek out new, challenging, or unusual work term or summer experiences? To what extent is the work term used to engage in new experiences or to test new skills and attitudes?

Students were rated higher for the first semester on this vector than on any other, and change is moderate. A dip at the third semester is puzzling. It may be this dip is a period of recovery or quiescence after the difficult adjustments of the first year. Perhaps the student wants constancy and safety for awhile to consolidate some of the new positions achieved.

In another somewhat similar institution Lois Murphy observed: "We are familiar with students who find the multitudinous change involved in leaving their home settings to come to Sarah Lawrence overstimulating, especially when the home setting is very different from what they find at college. The experience of overstimulation is increased by the multitude of choices that must be made and the degree of responsibility for planning one's own program, the lack of structured social groups, the need to find one's way socially as well as intellectually. In other words, some students feel themselves buffeted about by so many new currents and new experiences, new opportunities, demands, and challenges that it is hard to organize their lives." (7) These words certainly are apposite to the first year experiences of the Goddard student, and as a result the student may be content to be somewhat free from new challenges and from experiences which might create additional disequilibrium during his third semester.

Resourcefulness and Organization

The faculty described both the independent student and the purposeful one as resourceful and well organized. He is practical and able to work out intermediate steps to a goal. He knows when he needs help and how to get it; he is efficient, and knows how to make good use of the resources available to him. Four questions were used for this vector of development:

1. How freely does the student make use of a wide range of resources for his own learning?

2. How well does he make plans, follow them, or modify them consciously and judiciously and then follow them?

3. How well does he discover or develop new ways of approaching matters of concern to him? Of circumventing or overcoming obstacles that appear?

4. To what extent is he able to handle a variety of responsibilities and sustain good effort and performance in relation to all? How well is he able to avoid being overwhelmed at the end of the semester or at other times when several obligations seem to coincide?

Here again the picture is one of substantial change. Appearing for the first time is the Senior Division slump. There is a sharp drop of performance in the fifth semester after admission to the Senior Division. Once over the hurdle, the student comes down on the other side about where he was when he started. Fortunately, he must continue, and there is another higher hurdle looming ever closer, for which he rather quickly seeks altitude. This pattern recurs in the last two vectors of change with even greater force.

Full Involvement, Motivation, and Persistence

The faculty definitions of the independent student and of the purposeful student both gave a prominent place to motivation and persistence. The independent student, it was said, is motivated and working for his own satisfaction. He has the energy and determination to finish a job. The purposeful student is willing to tackle routine or difficult jobs congruent with his purposes and is resistant to obstacles. He continues in spite of mistakes or difficulties. He can sustain effort in the face of distractions and seeks out extra activities in addition to academic work which relate to his goal. The questions used were:

1. In general, how well motivated, persistent, and fully involved was this student?

2. What was the general level of effort reflected in preparation for classes, work on papers, and in relation to other kinds of responsibilities?

3. How consistent, steady, and regular was the student's output?

4. How great was the student's interest, enthusiasm, and intensity of

involvement with his work?

5. How good was his attendance in relation to the general expectations and the nature of the class?

6. What was his level of participation as compared to that which seems to be usual or satisfying for him?

As with Resourcefulness and Organization, there is fairly substantial change with a quite similar pattern. There is an even sharper drop after admission to the Senior Division; fifth semester ratings are substantially below those for either the third or fourth semester. They do remain higher than those for the first two semesters; hence, there is not complete regression. However, the recovery is so much stronger for the sixth semester that the students are about where they would have been had the rate of development continued uninterrupted.

Interdependence

Interdependence was the variable on which ratings reflected least change. Mature independence rests partly upon recognition of one's dependencies, upon recognizing that essentially one is involved with a network of interdependencies. The independent student therefore, the faculty said, is non-punitive, non-hostile, attuned to the whole, aware of his own resonances with it, and his own responsibilities to it. The questions they used were:

1. Is the student ready and able to work with others on community affairs such as recreation events, community government, house business, etc?

2. Does the student pull together with others well on work programs? To what extent is he conscious of his role in a broader work-program context, when such a relationship exists?

3. Does he seem to be aware of the relationship between his own behavior and community welfare in general?

4. Is the student tolerant of differences in behavior or in point of view on the part of other students or faculty members?

It is interesting to note that the sharpest change occurs in the third semester, and that fourth semester ratings are nearly as high as those for the seventh semester. This general pattern is consistent with our observation that second year students become heavily involved in community activities, and that the involvement does not increase much beyond; ratings for the sixth and seventh semester are about the same as those for the third and fourth. Here again, as with the previous two vectors of change, there is a conspicuous decline during the fifth semester.

Discussion

There seem to be at least two general patterns of change and additional secondary patterns depending upon how far one thinks the data can be pushed. There is the rather even and regular pattern of change for Goal Directedness and Personal Stability and Integration, and the pattern with the fifth semester slump as reflected for Resourcefulness and Organization, Full Involvement Motivation and Persistence, and Interdependence. Within each of these two clusters further distinctions are possible. In the case of Goal Directedness, the fourth semester seems to carry particular weight, and in the case of Personal Stability and Integration it is the first and seventh semesters.

Within the other cluster, change in Resourcefulness and Organization, and Interdependence occurs during the first two years, while it is during the Senior Division semesters that greatest total gain in Involvement, Motivation and Persistence occurs, even though this vector shares with the other two a sharp regression during the fifth semester. Venturing, with its sharp decline during the third semester remains in a class by itself, although it shares with Interdependence the picture of rather minimal change during the Senior Division years.

In general, most change seems to occur during the first two years. However, more detailed examination indicates that with Goal Directedness and Personal Stability, change is quite evenly dispersed over the four years, while Interdependence and Venturing change most during the first two, and Involvement, Motivation, and Persistence, change most during the last two.

Thus the question, "Do our students change while here, and is that change congruent with our objectives?" can be answered affirmatively, and the patterns of various vectors of change have also been susceptible to description. The validity of generalizing these patterns to other institutions has yet to be demonstrated and is beyond the scope of this study. If differences in institutions produce differences in development, then some patterns would be unique (those relating to the Junior-Senior Division arrangement would be likely candidates, for example.)

But what of the third question? Are there developmental principles which apply to student change in college? Four seem to receive support, albeit tentative:

First, development occurs according to generalizable sequences. That is to say, when a group of relatively similar persons undergo relatively similar experiences, observed over a period of time, change occurs according to recognizable patterns, patterns which differ depending upon the particular kind of change under consideration. Erik Erikson says: "Whenever we try to understand growth it is well to remember the epigenetic principle derived from the growth of the organism in utero. Somewhat generalized this principle states that anything that grows has a ground plan, and that out of this ground plan the parts arise, each part having its time of special ascendancy, until all parts have arisen to form a functioning whole.... it is important to realize that in the sequence of his most personal experiences the healthy child, given a reasonable amount of guidance, can be trusted to obey the inner laws of development, laws which create a succession of potentialities for significant interaction with those who tend him." (5) The distinctive patterns discovered above suggest that this principle still holds during the college years.

Second, development occurs through sequences of differentiation and integration. As Nevitt Sanford has formulated it, "A high level of development in personality is characterized chiefly by complexity and by wholeness. It is expressed in a high degree of differentiation, that is, a large number of different parts having different and specialized functions, and in a high degree of integration, that is, a state of affairs in which communication among parts is great enough so that the different parts may, without losing their essential identity, become highly organized into larger wholes in order to serve the larger purposes of the persons....This highly developed structure has a fundamental stability which is expressed in consistency of behavior over time....But the structure is not fixed once and for all, nor is the consistence of behavior absolute; the highly developed individual is always open to new experience, and capable of further learning; his stability is fundamental in the sense that he can go on developing while remaining essentially himself." (8) Such development may occur in a step-like process where increasing differentiation is accomplished by an acceleration in behavioral change, followed by development of integration at a higher level, during which deceleration of change in external behavior occurs. The patterns for Goal Directedness and for Personal Stability and Integration conform to this principle quite well.

Third, development is congruent rather than compensatory. Terman's longitudinal study of gifted persons (10) well documented this principle, and it is supported further here. Change occurs in all vectors, and not in some at the expense of the others. Although the patterns of change differ, the relationships among them do not suggest that lack of change in one area is compensated by greater change in another. When the figures are laid one above the other, the picture is more of a single cord proceeding upward with interweaving strands, than of two separate cords proceeding in opposite directions.

Fourth, development decreases as relevant conditions become more constant. Thus, the approach of entrance to the Senior Division provokes acceleration in Goal Directedness, and with assimilation of the new experiences of the first semester and in anticipation of graduation Personal Stability and Integration move to higher levels; plateaus and regression follow admission to the Senior Division and entrance into the comfortable Junior Year. Freedman makes a similar observation of Vassar alumni, observing that "the increased stability and well-being of alumnae as compared to seniors is primarily a product of the less rigorous lives of the former, the lessened intensity of the demands made upon them." (6)

These principles clearly have relevance for institutions planning innovation and experimentation, or institutions facing decisions concerning current practice. If the dimensions of development for students in an institution can be identified and patterns of change described, then questions concerning the nature of experiences to be introduced and the timing and location of their introduction can be answered more soundly. The existence of plateaus and points of regression suggest periods during the student's experience where enrichment of conditions or additional stimuli relevant to that vector of change might be helpful. Consequently, the fifth semester slump in three variables at Goddard College suggests the need for some attention to what is asked of students during this time, some attention to the differences in student experience during this period as opposed to others. And finally, as long as the strands of development generally move together, new programs, new conditions of living, new experiences can be added to foster change in some areas, without great risk of diminished development in others.

Finally, numerous questions remain which these studies only raise more sharply. These discussions concern students who graduated, who successfully completed a four year experience. What is the picture for those who left before graduating? What kinds of development occur among the drop-outs who are at the institution for varying lengths of time? And more generally, what really accounts for these differences in pattern and timing? What elements of the college program operate significantly in these various areas? What in the lives of students affects them causing these different kinds of development to occur? And how permanent are these changes? How much have the changed behaviors, reflected in the records, resulted in more enduring changes in the individuals themselves, changes which will be sustained in future contexts more benign or malignant?

References

1. Chickering, A.W., "Development of Independence," *Journal of Higher Education*, Vol. 35, No. 1., Jan. 1964.

2. Chickering, A.W., *Development of Purpose*, Plainfield, Vt.

3. Beecher, G., Chickering, A.W., Hamlin, W.G., and Pitkin, R.S. *An Experiment in College Curriculum Organization*, Plainfield, Vt., 1966.

4. Cooley, W.W. and P.R. Lohnes, *Multivariate Procedures for the Behavioral Sciences*, John Wiley and Sons, Inc., 1962, pp.61-63.

5. Erikson, E., "Identity and the Life Cycle," *Psychological Issues*, I (1959), 52.

6. Freedman, M.B., "Studies of College Alumni," ed. N. Sanford, *The American College*, John Wiley and Sons, Inc., 1962. pp. 864-65.

7. Murphy, L.B. and Raushenbush, E., *Achievement in the College Years*, New York, 1960.

8. Sanford, N., "Developmental Status of the Entering Freshmen," ed. N Sanford, *The American College*, John Wiley and Sons, Inc., 1962, p. 257.

9. Shoben, E.J., *Students, Stress, and the College Experience,* Washington, D.C., 1966.

10. Terman, L.M. and Oden, Melita, *The Gifted Child Grows Up*, Stanford, Calif.

11. U.S. Office of Education and The Union for Research and Experimentation in Higher Education, *Conference on Innovation in Higher Education*, 1966.

YEAR ROUND EDUCATION AND THE INSTRUCTIONAL PROCESS

Irma T. Halfter
Director, University Evaluations
De Paul University

Many colleges and universities are endorsing the principle of acceleration through calendar change for the full-time students in their institutions. The quality of such accelerated education in vigorously debated by faculties and administrations, but not its assumed consequence: that per capita cost of education will be less. Schoenfeld and Schmitz, in their thorough and concise study (1964) of the many forms of accelerated or year round education, conclude: 1) Acceleration, now thoroughly researched and evaluated, is widely accepted; yet the experiments and studies have focused primarily on the elite fractions of the student body; 2) prevention of decline in quality of education through acceleration depends on obtaining "regular," full-time students to attend and teachers to instruct (in short, the academic calendar design does not exist in a vacuum, and student enrollment must be approximately equal in each term, otherwise operational and capital savings fail to materialize); 3) institutions which have substantial summer enrollments for special audiences must evaluate the academic and financial consequences if the university adds or switches to a target enrollment of regular, degree students.

Acceleration

Educational change is seldom confined to elite students or elite institutions. Urban, commuting universities (such as De Paul University) propose calendar-curriculum change to year round study for their working-class degree students, who are at least partially or substantially self-supporting. Such institutions also have large, special audience enrollments in the summer. To suggest consequences for curricular-calendar change institutional research in such institutions may need to establish, before such changes in the calendar, what resources of the university full-time day students used (when they used them, and for how long) and whether graduating student behavior differed from non-graduating student behavior. One approach is to establish Modal Patterns of Attendance on the assumption this information may provide a base for decision-making for two categories of problems: the effect of year-long education on finance and work-experience for the described students, and the probability of year-long attendance occurring, to keep at least level operating and capital cost of the institution. Modal Patterns of Attendance should indicate, for example, whether students had utilized off-time offerings (e.g. summer sessions or evening study), how extensively this occurred, whether they accelerated or decelerated study, and whether attendance was consecutive or non-consecutive.

Situation

De Paul University, with a two-semester day calendar, will change to a four-quarter calendar, day and evening. The calendar change in September, 1967, implements a total curricular revision on a university basis as well as the formation of a new college and the restructuring of the relationships of all other administrative units and colleges, day and evening, to the new college. Many in the university had assumed: 1) Regular, full-time day students were earning their degrees in eight semesters of consecutive day study; 2) Students who did not earn their degrees in eight consecutive semesters were non-graduates and "De Paul drop-outs."

Procedure

A longitudinal study of the records of 667 full-time day students entering September 1961, and still in progress, is establishing Modal Patterns of Attendance by college of enrollment on entrance, at any subsequent time, by sex, and in two categories: graduates and non-graduates.

Some Findings

1) The permutations and combinations of patterns of attendance were so numerous that principal patterns had to be selected for study. 2) Whereas with graduates there are six major patterns of attendance, with non-graduates (at the point of time of the study) there are eighteen major patterns (CF. selected charts). 3) The university (and all sub-group) graduates did not typically earn their degrees by the semester-only pattern of study. 4) All categories of university graduates utilized summer sessions and/or other times, although there were differences by college and sex. 5) Summer sessions were utilized by non-graduates if the non-graduates completed four semesters of study. 6) Summer sessions were generally utilized by university graduates more than once, i.e. sixty-five per cent attended two or more summer sessions. 7) Consecutive or non-consecutive semester patterns of attendance varied with graduates and non-graduates, with the graduates in four years having predominantly a consecutive-semester pattern of attendance--with or without summer school and other times. 8) A university graduate had slightly more than a fifty-fifty chance of being graduated with the prescribed number of hours. (This incidental finding has provoked much "soul searching" by the departments and the university of why students are graduating with more than the required hours. Academic incompetence is not the reason. Students in Commerce, for example, had a better chance of graduating within the hour-requirement than in Liberal Arts.) 9) De Paul was experiencing a "drop in" of students rather than a "drop out" of students. The study of individual patterns of attendance confirmed a previous tentative finding: the graduating rate of entering full-time day students in 1961 is now (February, 1967) fifty per cent, as contrasted with a four-year graduating rate (June, 1965) of thirty-four per cent. With thirteen per cent of 1961 non-graduates now in attendance, the graduating rate of the original 1961 group may be expected to reach fifty-five per cent. In short, decelerators were substantial in number. 10) The eventual graduating rate of women will approximate that of men--forty-seven per cent versus fifty per cent.

Conclusions

The utilization of summer sessions and of other times for degree study would seem to indicate year round education will not be an altogether strange notion for De Paul students. On the other hand, the findings do not warrant stating that the attendance of these students would alone support a summer session at De Paul. The report does not indicate why this use of summer sessions or other times occurred. The relationship between year round education and year-round financing of education by working students is not established. For the designers of the new academic curricula and methods and modes of learning, however, some acceptable implications seem to be: The proposed sequential planning of courses and their being offered only in the autumn-winter-spring quarters may adversely affect students' flexibility of choice of time, place, and kinds of educational experiences. The pacing of the new interdepartmental offerings and restructuring of old, departmental offerings, in relation to the new divisional offerings, may be affected--to say nothing of increasing the problem of equating course credits over time and of the inevitable changes that occur with experimental designs.

REFERENCE

Schoenfeld, Clarence A., and Schmitz, Neil. Year-round Education: Its Problems and Prospects from Kindergarten to College. Madison, Wisconsin, 1964.

PREDICTABILITY OF COLLEGE GRADES:
DIFFERENCES BETWEEN TWO APPARENTLY LIKE COLLEGES

O. W. Hascall
Regional Director
American College Testing Program

The accuracy with which freshman grades can be predicted is greater for some colleges than for others. Why more precise predictions can be made for one institution than for another is usually vague and not often understood. This study is an attempt to pinpoint some of the factors which may account for a wide disparity in the accuracy of freshman college grade predictions at two colleges.

Method

To focus on this problem, the characteristics and dynamics of two colleges with many outward similarities were studied. Both colleges bear the nickname "School of Mines" in their respective states; both offer degrees in mining engineering, related engineering fields, and in non-related engineering disciplines; and both are approximately the same size. Furthermore, both institutions have participated in the ACT Research and Class Profile Services. For our purposes, the colleges are identified as "Western Tech" and "Midwestern Tech."

Table 1 presents data which explain why these two colleges were chosen for this study. On the basis of identical variables, the optimum correlation with first year college grades is relatively low at Western Tech but quite high at Midwestern Tech.

Table 1. Correlations Between Certain Variables and First Year Overall Grade-Point Averages at Two Colleges

Variables	Correlations					
	Western Tech			Midwestern Tech		
ACT scores						
English	.120	T-Index .205	TH-Index .373	.368	T-Index .569	TH-Index .694
Mathematics	.132			.517		
Social studies	.197			.440		
Natural science	.140			.345		
H. S. Grades						
English	.339	H-Index .391		.515	H-Index .639	
Mathematics	.239			.519		
Social studies	.179			.452		
Natural science	.336			.512		

In an attempt to discover some of the factors which may account for the differences in these correlations, several sets of information were examined. The ACT Research Reports for the two institutions contain the data from which the information in Table 1 was extracted. In addition to these kinds of correlations, the Reports contain correlations between eight variables and grades in certain subjects. Both colleges used grades in freshman English, freshman mathematics, and freshman chemistry as criteria. Western Tech also used grades in "History" and "History of Civilization" as criteria, while Midwestern studied the relationship between the eight variables and "engineering Graphics." The Research Reports also contain frequency distributions, means, standard deviations, expectancy tables, and other data which are by-products of the correlation statistics.

Another source of information was the ACT Class Profile Report for each of these colleges. That Report describes the characteristics of enrolled freshmen according to answered questions on the Student Profile Section of the ACT battery at the time they took the examination. In that Section, the student was asked to answer questions about his plans for college and about his personal and family background. The Class Profile Report describes the colleges' freshman classes in terms of the proportions of students who answered these questions in a given way.

Information about the way in which final examinations in certain courses were prepared and administered, grading practices in remedial courses, and proportions of students in certain curriculums was obtained by means of a questionnaire which was completed by administrators at the two colleges. Additional information about course requirements for freshmen was obtained from the catalogs of the two institutions.

Results

Tables 2 and 3 show the general nature of the courses which freshmen at the two colleges pursue. While there are slight differences, the general pattern of courses are quite similar at the two colleges.

Table 2. Required Freshman Courses in Engineering Curriculums

Western Tech	Midwestern Tech
Chemistry 10 hrs.	Chemistry 8 hrs.
Engr. Graphics 4 hrs.	Engr. Graphics 4 hrs.
English 6 hrs.	English 6 hrs.
Anal. Geom. and Calc. 10 hrs.	Anal. Geom. and Calc. 8 hrs.
Phys. Educ. 2 hrs.	Phys. Educ. 2 hrs.
	Mil. Sci. 2 hrs.
Electives 6 hrs.	Electives 6 hrs.

Table 3. Required Freshman Courses in Most Non-Engineering Curriculums

Western Tech	Midwestern Tech
English 6 hrs.	English 6 hrs.
Modern History 10 hrs.	Soc. Sci.
or	or
Hist. of Civ. 6 hrs.	Humanity Elec. 6 hrs.
Mathematics 8 to 10 hrs.	Anal. Geom. and Calc. 8 hrs.
or	
Chemistry 10 hrs.	Chemistry 8 hrs.
or	
Geology 6 hrs.	
Phys. Educ. 2 hrs.	Phys. Educ. 2 hrs.
	Mil. Sci. 2 hrs.
Elective 0-6 hrs.	Elective 4 hrs.

Table 4 presents a comparison between the two colleges on the basis of certain student characteristics. Generally the characteristics of freshmen on both campuses are quite similar. However, a few striking dissimilarities are noteworthy. At Western Tech, only about one-fourth of the students are pursing an engineering curriculum while about three-fourths of Midwestern Tech freshmen are enrolled in an engineering program. That 82 per cent of Western Tech's freshmen plan to work, compared with 59 per cent of Midwestern Tech's freshmen, probably indicates another difference in the characteristics of the two student bodies. Some further differences are observed in the "type of home community" from which the freshmen came.

Table 4. Demographic Characteristics

	Western Tech	Midwestern Tech
Freshman Enrollment		
Men	257	309
Women	52	14
	309	323
Number in Engr. Curriculum	79 (26%)	250 apx (77%)
Number in Non-Engr. Curriculum	230 (74%)	123 apx (23%)
Educational Plans		
Per cent planning for bachelor's degree	51%	48%
Per cent planning for some graduate work	30%	44%
Other	19%	8%
Housing Plans		
College dorm	23%	59%
At home	69%	27%
Other	8%	14%
Per Cent Planning to Work Part Time	82%	59%
Most Important Goal in Attending College		
To develop mind and intellectual abilities	31%	30%
To secure vocational or professional training	53%	56%
Other	16%	14%
Type of Home Community		
Farm or open country	12%	40%
City or suburb under 100,000 population	80%	52%
Other	8%	8%
Age at Enrollment		
17 or younger	11%	16%
18	53%	64%
19	18%	4%
20 or older	18%	16%
Marital and Dating Status		
Married	5%	7%
Engaged, going steady, or date same person	27%	30%
Date more than one person	46%	41%
Do not date at all	16%	16%
Other	6%	6%
Estimated Family Income		
Less than $5,000	17%	19%
$5,000-7,499	38%	28%
$7,500-9,999	11%	15%
$10,000 and over	11%	18%
Do not know	20%	15%
Considered confidential	3%	5%

The figures in Table 5 present a comparison of the two groups of students according to their academic ability (as measured by the ACT composite score and the average of four self-reported high school grades) and the freshman college grades they received. The students at Midwestern Tech have consistently higher mean high school grade averages than those at Western Tech. However, in three areas (English, mathematics, and history) Western Tech's freshmen earned a higher college grade-point average than Western Tech students. The highest mean grade-point average in Table 5 is 2.57 for Western Tech's freshmen who were enrolled in "History" and "Civilization" courses. Note that this mean is at the 98th percentile when compared to the mean grade-point averages in other social study courses at other colleges which have participated in recent ACT Research Services.

Table 5. Academic Characteristics

Students Enrolled In	Western Tech			Midwestern Tech		
	ACT Composite Score	Aver. of 4 h.s. grades	College g.p.a.*	ACT Composite Score	Aver. of 4 h.s. grades	College g.p.a.*
English						
Mean	21.4	2.55	2.19	25.1	2.95	2.01
S.D.	4.58	.72	1.11	3.32	.66	1.00
%-ile of Mean**	80	45	64	99	93	36
Mathematics						
Mean	22.2	2.66	2.16	25.3	2.95	1.87
S.D.	4.47	.80	1.24	3.26	.67	1.18
%-ile of Mean**	91	62	71	99	93	28
History and Civiliz.						
Mean	21.3	2.58	2.57			
S.D.	4.82	.74	1.03			
%-ile of Mean**	79	47	98			
Engr. Graphics						
Mean				25.2	2.96	2.07
S.D.				3.34	.66	1.01
Chemistry						
Mean	21.1	2.59	2.27	25.3	2.97	2.42
S.D.	4.90	.77	1.42	3.23	.64	.96
%-ile of Mean**	76	50	88	99	95	97
Overall						
Mean	21.7	2.58	2.15	25.1	2.95	2.23
S.D.	4.19	.69	1.04	3.32	.66	.80
%-ile of Mean**	84	49	51	99	93	68

* Refers to English g.p.a. in English area, mathematics g.p.a. in math area, etc.
** Refers to the percentile rank of this mean compared to the means of all colleges in recent ACT Research Services.

The information in Tables 6 and 7 is taken from the responses given on the questionnaire. While the freshmen at Midwestern Tech who enroll in first semester English take the same final examination regardless of which instructor teaches the course, Western Tech's freshmen take different final examinations in English, depending upon which instructor teaches the class. The information in Table 7 reveals further differences in the grading practices of the English classes at the two colleges and differences in the proportion of freshmen enrolled in the remedial English sections.

Table 6. Freshman Courses in Which a Common (Departmental) Final Examination Was Administered

Western Tech	Midwestern Tech
Mathematics	Mathematics
Chemistry	Chemistry
Geology	English (fall semester only)
History	
Hist. of Civilization	

Table 7. Selected Responses to Items on the Questionnaire

	Western Tech	Midwestern Tech
In what disciplines are so-called "remedial" or slow-moving courses offered?	Mathematics (007,008,005) English Comp. (099)	Mathematics (91) English (110)
Are there about the same number of A's and B's given in these "remedial" classes as in the "regular" classes?	Yes	No. Less A-B grades; more D-F
Additional Comments.	This past year, forty percent of all entering freshmen took the "slow" English course.	English 110 (remedial) --in which about ten percent of our freshmen have been enrolled --will be dropped next year.

Discussion

The results of this study reveal some differences between these two technological colleges. Whether or not any one or several of these differences might account for the difference in predictability of freshman grades may be examined by proposing the following hypothesis: Freshman grades at Western Tech may be more a function of the curriculum and courses in which the student enrolls than a function of his attributes as measured by the eight variables which have been correlated with college grades. At Midwestern Tech, these functions are controlled in such a way that they have little or no effect on the students' overall grade-point averages.

This thesis may be examined and illustrated by presenting the records of two hypothetical students (Table 8). On the basis of ACT scores and high school grades, John is portrayed as a student for whom a lower overall freshman college grade-point average would be predicted than for Frank. However, according to the theory which has been proposed, it is entirely possible for John to achieve higher freshman grades at Western Tech than Frank. First (as illustrated in Table 8), John chose to enter the

Table 8. Hypothetical Records of Two Students

			Possible First Year College Grades			
			Western Tech		Midwestern Tech	
	ACT Scores					
John (chooses non-engr. curriculum)	English	14	English (rem.)	B	English (rem.)	D
	Math	15	Math (rem.)	B	Math (rem.)	F
	Soc. St.	12	History	C	Chemistry	D
	N. Sci.	13	Phys. Ed.	A	Humanity	C
			Elective	B	Mil. Sci.	C
					Phys. Ed.	A
			Overall g.p.a.	2.5	Overall g.p.a.	1.2
	H. S. Grades					
	English	C				
	Math	B				
	Soc. St.	C				
	N. Sci.	C				
	ACT Scores					
Frank (chooses engineering curriculum)	English	22	Chemistry	C	Chemistry	C
	Math	24	Graphics	B	Graphics	B
	Soc. St.	20	Engl. (reg.)	D	Engl. (reg.)	C
	N. Sci.	21	Math (reg.)	C	Math (reg.)	C
			Phys. Ed.	B	Mil. Sci.	B
					Phys. Ed.	B
			Overall g.p.a.	2.2	Overall g.p.a.	2.5
	H. S. Grades					
	English	B				
	Math	A				
	Soc. St.	B				
	N. Sci.	A				

non-engineering curriculum at Western Tech which provided him an entry into the high-grading history class, an opportunity not afforded Frank because he chose the engineering program. Further, John enrolled in the remedial English and mathematics courses (because of weaknesses in these areas) but was able to earn B's in both of these classes because of the grading practices at Western Tech. Frank, on the other hand, enrolled in the more rigorous English and mathematics classes with a "tough" grading English instructor; therefore, he made only a D in English. If these two students had enrolled at Midwestern Tech, John would probably have made a much lower overall grade point average than Frank because students in both engineering and non-engineering programs take equally rigorous curriculums and because a disproprtionate number of D's and F's are given in the remedial English and mathematics courses at Midwestern Tech.

This explanation is only a theory. The next step in this study will be to examine the individual records of freshmen included in the study to see if the actual records do, in fact, resemble the hypothetical cases of John and Frank.

The results of this study have also been examined in the light of four common explanations for high correlations between variables and college grades as follows:

1. Unusually high correlations are often obtained when there is great variability in the academic potential of the students included in the sample and/or when there is

great variability in the college grades given. This explanation for the difference between the predictability of college grades at Western Tech and Midwestern Tech must be rejected. Western Tech's students, for which the lower correlation was obtained, have considerably more variability than Midwestern Tech's students. At Western Tech the standard deviation for the ACT composite score was 4.19, compared to 3.32 for Midwestern Tech; the standard deviation of the average of the four high school grades was .69 for Western Tech, compared with .66 for Midwestern Tech; and the standard deviation for the overall college grade-point average at Western Tech was 1.04, compared to .80 at Midwestern Tech (Table 5).

2. Unusually high correlations are often obtained when the college courses are particularly homogeneous; that is, when the freshmen are generally taking the same or very similar courses as opposed to a wide variety of subjects. This statement may be accepted as a partial explanation for the difference between the predictability of grades at the two colleges. The information in Tables 2, 3, and 4 indicates that the students at Midwestern Tech are pursuing a more common freshman curriculum than are the students at Western Tech.

3. Higher correlations may be expected when the environmental conditions of the students are standardized; that is, when all freshmen live on campus, none are involved with part-time work, etc. There may be some merit to this explanation when it is noted from the data in Table 4 that a greater proportion of Midwestern Tech's students do plan to live in college dormitories, and fewer plan to work part-time than do Western Tech's students.

4. When the sample of students studied contains a high proportion of those who are extraordinarily oriented toward academic goals, high correlations may be expected. Some of the results of this study indicate that this explanation may account, in part, for the higher correlations that were obtained at Midwestern Tech. A larger proportion of those students do plan to do some graduate study (Table 4); and their academic potential as evidenced by mean ACT scores and mean averages of four high school grades are significantly higher than students at Western Tech. However, when these two groups of students were asked to indicate their most important goal in attending college, the responses for the two groups were very similar (Table 4).

164/165

STUDENT PERCEPTION OF GRADING PRACTICES AT THE UNIVERSITY OF WASHINGTON

James K. Morishima
Acting Director, Office of Institutional Educational Research
University of Washington

Each winter the Registrar's Office of the University of Washington distributes a report which summarizes the grade distributions in various courses offered at the University during the fall quarter. Informal analyses of these reports reveal disparate grade distributions between sections of the "same" course.

Armed with the foregoing information, the Associated Students of the University of Washington (ASUW) Educational Affairs Commission (to which the author was a faculty representative) requested a study by the Office of Institutional Educational Research.

The results of that study (Morishima, 1963) indicated that one-half the students at the University of Washington had positive reactions to the general grading procedure at the University. The study reported herein is an extension of the earlier study.

Method:

As an adjunct to a study of examination schedules and student behavior (Morishima, 1967), the Office of Institutional Educational Research administered a questionnaire designed to yield data relevant to student satisfaction with certain aspects of the grading practices of various courses at the University of Washington. It was hoped that the experience gained from a previous report (Morishima, 1963) and the open-ended segment of the questionnaire would shed light on the degree of satisfaction felt by students with the grading systems utilized in their various courses.

It was well recognized at the outset that the subject of "grading practices" was a complex beast with a seemingly innocuous title, and that students would be answering the questionnaire with various grading practices in mind. In an effort to provide some anchor, students were asked to fill out one questionnaire for each course they were taking.

Three hundred subjects (125 Ss during the first week of instruction, spring, 1966, and the 175 Ss who participated in an earlier study (Morishima, 1967) during the last week of instruction, spring, 1966) were randomly selected, and a trained interviewer contacted them at a mutually convenient time and place. Each S was asked to fill out one questionnaire for each course he was taking (excluding physical education and R.O.T.C.). Upon completion of the questionnaires, questions regarding the various courses, instructors, etc., were asked by the interviewer in a highly unstructured interview format.

Since there were no reliable differences between the two groups of students, the results have been combined in most of the analyses. Furthermore, the comparable questions from the earlier survey (Morishima, 1963) were not significantly different from the present survey.

Table 1 summarizes the responses to the question, "What is your opinion of the system for giving final grades in this course? Do you think it is a fair system or an unfair system for determining how much you have learned? I think the grading system for this course is: (check one)."

Table 1. Opinion of Grading System at University of Washington

Category	Male		Female		Total	
	N	%	N	%	N	%
Very fair	15	02.7	14	03.3	29	03.0
Quite fair	93	17.0	93	21.7	186	19.0
Pretty fair	179	32.7	136	31.7	315	32.2
Total positive	287	52.4	243	56.7	530	54.2
Somewhat fair	107	19.5	93	21.7	200	20.5
Somewhat unfair	111	20.3	78	18.2	189	19.3
Total "neutral"	218	39.8	171	39.9	389	39.8
Pretty unfair	25	04.6	10	02.3	35	03.6
Quite unfair	13	02.4	4	00.9	17	01.7
Very unfair	5	00.8	1	00.2	6	00.7
Total negative	43	07.8	15	03.4	58	06.0
Grand Total	548	100.0	429	100.0	977	100.0

Table 1 indicates that over one-half the courses taken by the students in the sample were seen as having a fair means of arriving at final grades.

When these data are further subdivided on the basis of courses giving most frequent (MF) or most hours (MH) of examination vs. courses giving least frequent (LF) or least hours (LH) of examination, students were found to rate the former as "fairer" than the latter ($P<.01$). In short, MF and/or MH courses were judged significantly more positively than LF and/or LH courses.

This result may be coupled with the data summarized in Table 2. Table 2 presents the results of a content analysis for MF or MH courses vs. LF or LH courses for the question: "What changes would you make in either the system of giving course grades or the mid-term and final examinations in order to make these fairer ways of determining how much you have learned in this course?"

Table 2. Suggestions for Changes

	Hours/Frequency of Examinations					
	Most		Least		Total	
Category	N*	%	N*	%	N*	%
More exams	5	01.6	94	19.9	99	12.4
More refined grading system	23	07.2	83	17.6	106	13.4
Make former exams available	44	13.7	60	12.7	104	13.1
More objective items	76	23.7	83	17.6	159	20.1
More subjective items	84	26.2	87	18.4	171	21.6
Other	35	10.9	42	08.9	77	09.7
No comment	54	16.7	23	04.9	77	09.7
Total	321	100.0	472	100.0	793	100.0

*Note: Totals may exceed 300, since some students volunteered more than one suggestion.

Two questions relate to our concern about the relationship between examination schedules and student suggestions for system changes. First, students who are not examined frequently request more examinations. Second, these same students want a more refined grading system. Both differences between "most" and "least" are significant at less than the .01 level.

Another interesting difference between the two groups appears. The LF and LH questionnaires contain more comments than MF and MH questionnaires. Note that there are no differences between the groups on requests for more objective test items, more subjective test items, or "others." There is a slight, though nonsignificant, tendency for students who are infrequently tested to request that past exams be made available in the library, for example.

It is of interest to note that when the residence of the students is taken into consideration there is a significant difference between students from each of the groups (fraternity and sorority, residence halls, and "commuters") in the frequency with which they requested: (1) that former exams be made available ($P = .02$), and (2) that more subjective test items be given ($P = .04$). (Moreover, Greeks requested objective items in tests slightly more frequently than did non-Greeks.) In other words, a few Greeks, slightly more "dormies," and even more commuters requested that past exams be placed on file somewhere, while commuters, dormies, and Greeks, in that order, requested more subjective items in examination. It appears as though we have not been suffering from pluralistic ignorance. Greeks do appear to have readier access to test files than do "dormies," who, in turn, have access to more test files than do the commuters. The foregoing statement is, however, an inference drawn from the data, since we did not ask whether students had access to test files.

It is of further interest to note that only fifty-eight of the questionnaires (see Table 1) revealed that the means of arriving at final grades was thought to be unfair. Yet Table 2 reveals that students had a great many suggestions for changes. Perhaps more refined analysis will shed some light on the reasons for this difference.

Discussion:

These data tend to indicate that most students at the University of Washington are satisfied with the present grading system. This result replicates in a different form reaction to Guthrie's (Wilson, 1963) proposal to enact a finer grading system at the University. Guthrie found opposition from three quarters: (1) faculty who complained that assigning students to one of five categories was difficult enough; (2) students who felt that a finer discrimination coupled with the gross achievement measures in use would be unfair to borderline students, i.e., students with very minimal GPA's; and (3) the administration, i.e., the Registrar's Office, which felt that the grading system would cause more administrative red tape and would make the University's grading system unlike that employed by most colleges and universities.

The data we have tend to indicate that before a finer grading system is introduced, changes in the discrimination of level of achievement must be made. That is, it is not very meaningful to further subdivide the gross categories (A through E) when these subdivisions cannot be accurately assigned. There was little tendency in the interview material, however, for students to call for a grosser grading system, e.g., pass-fail.

One finding bears special mention. If it is true that commuters have little if any access to previous examinations, then the faculty of an institution (such as the University of Washington) which has a large commuter population should be encouraged to make "back tests" available to all students in the course to minimize giving some students advantages over others. Such a system may be especially critical in courses where students are examined twice--once at mid-term and once at finals. One instructor at the University of Washington who was accoustomed to giving just a mid-term and a final changed his system somewhat. He gave a sample examination in class for class discussion before the mid-term examination. Results compared with the performances of students enrolled in the same course the preceding quarter indicated that students performed better on the mid-term and final in the sample examination procedure. In effect, students had learned what to expect from the instructor.

Finally, it is suggested that further work be done on this problem, utilizing either larger numbers of respondents or blanket coverage in selected courses and/or sections.

References

Morishima, J.K. "Preliminary Study of Students' Conception of University of Washington Grading Practices," Office of Institutional Educational Research, University of Washington. Report No. IER-63-1, October 23, 1963.

"The Effects of Examination Scheduling on Student Behavior," Office of Institutional Educational Research, University of Washington. Report No. IER-111-1, January 27, 1967.

Wilson, W.R. (1963). "Reactions to Guthrie's Proposal for a Grading System Utilizing Finer Discriminations." Personal communication.

AN ANALYSIS OF GRADING PRACTICES

Dwain F. Petersen
Post-doctoral Intern, Office of Institutional Research
Michigan State University

The use of a grading system can be likened to the use of a taxonomy of educational objectives. According to Bloom, et al, (1956): "...the major task in setting up any kind of taxonomy is that of selecting appropriate symbols, giving them precise and usable definitions, and securing the consensus of the group which is to use them." The symbols of grading have been selected, and whether or not they are appropriate is presently irrelevant; the A through F grading system is the custom throughout higher education. The symbols do not, however, have precise definitions nor is there a consensus concerning how they are to be used.

Purpose

It is my purpose to present data related to current grading practices, focus on some specific variables, and to project possible courses of action directed at more consistent treatment of these variables. To achieve these purposes for the current scene, illustrative data from a computer-produced "Grading Practices Analysis" report will be analyzed from the standpoint of the following questions: (1) Are there differences in mean grades assigned among the several colleges of a university or among levels of courses taught? Within a college? Within a department? (2) In the several sections of large enrollment courses, is there variability among, or differences in, mean grades assigned when common final examinations rather than individual instructor-prepared examinations are utilized? (3) Are there differences in mean grades assigned to honors sections compared with all other courses enrolled in during the same term by the same students? (4) Is there a relationship between section size and mean grade assigned in a section? (5) What are the grading practices in remedial courses in which grades are assigned? (6) What are some alternative courses of action and what would be the results of such action? These questions are only a sampling of the questions which could be asked as Juola (1967) has presented a paper based on this same grading analysis report.

Over thirty years ago at the University of Chicago it was stated: "The lack of reliability of teachers' marks has been pointed out so frequently in educational literature as to need no further demonstration" (Reeves, Peik & Russell, 1933). One might expect that in the intervening years the problems of grading have been resolved. Yet it appears that little if any progress has been accomplished toward solving them. Most recently Juola (1967), after studying grading practices at Michigan State University, concluded: "...many of our college-level grading practices border on being chaotic. Historical precedence has often seemed to provide the only basis for our practices and even this historical basis has been highly distorted." It might be argued that effort directed at ameliorating the situation is wasted; but since grades continue to be important to students just as salary is important to the professor, we should attempt to remove as many inequities as possible. To retain a proper perspective, however, we need to remember that there is much more to teaching and learning than salary and grades.

The Grading Practices Analysis Report

The data presented in this paper are relevant to one institution, but the variables exist throughout higher education. They are probably the result of similar factors in the various institutions of higher education. To isolate problems of grading practices the Office of Institutional Research at Michigan State University has directed the development of a "Grading Practices Analysis" report by the Systems and Programming Section of Data Processing. The report is generated from the grade card tape in student number sequence. The computer produces an analysis grade card record.

This analysis is based on the assumption that the average grade earned by students registered in a given section in all of the other courses in which these same students are enrolled during the same term indicates the average ability in the given section. There are many other variables involved in grading, but there is no reason to believe that they are not self-compensating among the registrants in any given section. I have used the most sensitive measure that can be derived from this assumption, which is the DIFFERENCE IN MEAN GRADES. The DIFFERENCE IN MEAN GRADES is the difference between MEAN GRADE IN SECTION and MEAN GRADE IN OTHER COURSES. The MEAN GRADE IN SECTION is also used frequently and is simply the average grade in a given section.

Results

To investigate differences by colleges and levels the "Grading Practices Analysis" for Fall, 1965, was sampled, and an analysis of variance of the mean grades in courses was accomplished. Eight of the thirteen colleges were selected in which at least ten courses were represented at the following levels: 100-299 lower division courses, 300-499 upper division courses, and 800-999 graduate courses. Ten courses were randomly selected in each of the three levels of the eight colleges. Table 1 presents an analysis of variance of the mean grades in this sample. This analysis indicates that differences in assigned grades exist among colleges and levels, but similar differences failed to appear in the interaction between colleges and levels. In Table 2 are presented the average grades by college and level for the random sample. This analysis does not control for different ability levels; it merely communicates the varying means of assigned grades.

TABLE 1. Analysis of Variance for Sample Mean Grades at Michigan State University, Fall, 1965

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F	Table Value
Total	239	103.18	-	-	-
Colleges	7	6.69	.96	4.00*	2.73
Levels	2	38.12	19.06	79.41*	4.71
Colleges x Levels	14	6.54	.47	1.95	2.17
Within	216	51.83	.24	-	-

*Significant at the .01 level

TABLE 2. Sample Mean Grades at Michigan State University, Fall, 1965

	Colleges								
Course Levels	Agriculture	Arts & Letters	Business	Communication Arts	Education	Engineering	Natural Science	Social Science	Average
100-299 Low.Div.	2.66	2.67	2.59	2.52	2.63	2.60	2.32	2.45	2.56
300-499 Up.Div.	3.42	2.90	2.33	2.88	3.08	2.73	2.95	2.43	2.84
800-999 Grad.	3.78	3.46	3.15	3.57	3.32	3.69	3.62	3.45	3.50
Average	3.29	3.01	2.69	2.97	3.01	3.01	2.96	2.78	2.97

The difference in mean grades was used in a similar analysis to take advantage of the control for ability level which it affords. Table 3 presents an analysis of variance of the differences in mean grades for the sample. This analysis indicates that differences exist among colleges, but similar differences failed to appear among levels of courses when ability level is controlled. The interaction between colleges and levels was also significant. It is indicated, although not conclusively, that differences in mean grades do not vary much from one level to another. Table 4 presents the average difference in mean grades by colleges and levels for the random sample. Little consistency can be perceived by levels or by colleges except in Agriculture and Communication Arts.

TABLE 3. Analysis of Variance for Differences in Mean Grades at Michigan State University, Fall, 1965

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F	Table Value*
Total	239	62.20	-	-	-
Colleges	7	10.40	1.48	7.40*	2.05
Levels	2	.20	.10	.50	3.04
Colleges x Levels	14	7.98	.57	2.85*	1.74
Within	216	43.62	.20	-	-

*Significant at the .05 level

TABLE 4. Sample Differences in Mean Grades at Michigan State University, Fall, 1965

	Colleges								
Course Levels	Agriculture	Arts & Letters	Business	Communication Arts	Education	Engineering	Natural Science	Social Science	Average
100-299 Low. Div.	+.27	-.16	+.09	+.14	+.36	+.37	-.17	+.05	+.12
300-499 Up.Div.	+.85	+.13	-.30	+.41	+.46	-.03	+.10	-.14	+.18
800-999 Grad.	+.78	+.04	-.03	+.22	-.13	+.19	+.26	-.03	+.16
Average	+.63	.00	-.08	+.26	+.23	+.17	+.07	-.04	+.16

To illustrate the differences in mean grades assigned within a college, summary data for each department in the College of Arts and Letters are presented in Table 5. The College of Arts and Letters was selected for this illustration because it had among the smallest sample difference in mean grades at each level in Table 4, and its overall average difference in mean grades was near zero. Some of the departments in Table 5 had statistically significant differences which are probably of no practical significance. For example, Department E had a difference in mean grades of only .08 of a grade point. This means that approximately one out of thirteen students would be assigned one grade higher, in addition to the expected variability in grade assignment. On the other hand, in Department N one out of two students would be so treated.

TABLE 5. Summary Data for Department in the College of Arts and Letters at Michigan State University, Fall, 1965

Department	Number of Grades in Section(s)	Mean Grade in Section(s)	Number of Grades in Other Courses	Mean Grade in other Courses	Difference in Mean Grades	T-Test of Difference*
A	37	3.24	101	3.01	0.23	1.22
B	1938	2.79	6006	2.52	0.27	10.98*
C	201	2.51	630	2.96	-0.45	-4.68*
D	49	2.64	168	2.76	-0.12	-0.53
E	4276	2.57	13806	2.49	0.08	4.60*
F	1843	2.35	6523	2.61	-0.26	-9.48*
G	43	3.21	148	2.94	0.27	2.75
H	1046	2.39	3689	2.86	-0.47	-13.30*
I	4828	2.29	16172	2.47	-0.18	-10.67*
J	108	2.72	357	2.44	0.28	3.04*
K	25	2.97	88	2.81	0.16	1.08
L	63	3.23	164	3.18	0.05	0.22
M	134	3.07	485	2.74	0.33	4.43*
N	3010	3.25	15371	2.74	0.51	27.45
O	1360	2.51	4557	2.56	-0.05	-1.45
P	666	2.70	2380	2.42	0.28	8.67*
Q	38	3.13	134	2.42	0.71	3.92*
R	307	2.85	1115	2.75	0.10	2.25
S	1391	2.63	4730	2.54	0.09	4.18*

*Significant at the .01 level

To illustrate the differences in mean grades assigned within a department, a department with at least 1,000 students and an insignificant t-ratio of less than 1.96 was selected from the College of Arts and Letters. Department 0 in Table 5 met these criteria and each upper division section with at least 20 students (required for the use of the t-test) is analyzed in Table 6.

TABLE 6. Grading Practices Analysis of Upper Division Sections of Department 0 of the College of Arts and Letters at Michigan State University, Fall, 1965

Number of Grades in Section	Mean Grade in Section	Number of Grades in Other Courses	Mean Grade in Other Courses	Difference in Mean Grades	T-Test of Difference*
23	2.74	79	2.71	0.03	0.02
30	2.30	101	2.52	-0.22	-0.79
64	2.91	207	2.74	0.17	2.18
25	2.76	77	2.12	0.64	2.97*
36	2.17	122	2.27	-0.10	-0.05
38	2.47	129	2.65	-0.18	-0.74
30	2.20	93	2.47	-0.27	-1.00
67	3.10	235	2.67	0.43	3.14*
26	3.15	76	2.56	0.59	5.10*
99	2.73	315	2.49	0.24	2.32
21	2.62	51	2.86	-0.24	-0.59
35	3.00	87	2.57	0.43	2.26

*Significant at the .01 level

Upper division sections were selected to further reduce the expected variability for this illustration. It is demonstrated in Table 6 that in a department in which one might expect normal grading practices, there are still sections with significantly different grades assigned to its registrants than might be expected.

To investigate the question about common final examinations, the courses in which common finals were administered to at least ten sections of a course were matched with courses of a similar level and discipline where common finals were not administered. The variance of the difference in mean grades was calculated for the groups of sections and a non-parametic technique, Wilcoxin's Matched-Pairs Signed-Ranks Test (Downie & Heath, 1965), was used to investigate the degree of variability. In Table 7 are presented the results of this procedure. The T of 10 indicates that there is a significant difference in the variability of grading practices where common final examinations were and were not used. There was less variability in courses in which common final examinations were used. To determine whether there were differences in the mean grades assigned to this sample, the same procedure described above was applied to the average of the differences in mean grades assigned to the courses rather than the variance. In Table 8 are presented the findings. A T of 39 indicates, although not conclusively, that average grades assigned are no different in courses where common finals were and were not administered.

TABLE 7. Matched-Pairs Signed-Ranks Test of Variance in Grading Practices in Courses in Which Common Final Examinations Were and Were Not Administered

Discipline	Level	Common Finals Variance	Othen Than Common Finals Variance	Difference	Absolute Rank of Difference	R(+)	R(-)
Language	L.D.	.0513	.1334	.0821	10	10	
Language	L.D.	.0932	.0788	-.0144	3		3
Prof.	L.D.	.0439	.1629	.1190	14	14	
Prof.	U.D.	.2715	.3157	.0442	7	7	
Science	L.D.	.0154	.0092	-.0062	1		1
Science	L.D.	.0557	.1633	.1076	13	13	
Science	L.D.	.0858	.2710	.1852	15	15	
Science	U.D.	.0289	.0572	.0283	5	5	
Science	U.D.	.0669	.0903	.0234	4	4	
Science	L.D.	.1543	.2280	.0737	8	8	
Science	L.D.	.0857	.1745	.0888	11	11	
Science	L.D.	.1075	.0690	-.0385	6		6
Science	L.D.	.0348	.3129	.2781	16	16	
Science	L.D.	.0543	.1303	.0760	9	9	
Science	L.D.	.0274	.1190	.0916	12	12	
Soc.Sci.	L.D.	.0544	.0641	.0097	2	2	
Median		.0550	.1318	.0748	Total	126 T =	10

A table value of 20 or less required for a two-tailed test at the .01 level (Downie & Heath, 1965)

To investigate the question regarding mean grades assigned to honors sections, all of the honors sections were grouped together by college. Correlated t-tests, comparing mean grades assigned to honors sections with mean grades assigned to the same students in all other courses during the same term, were calculated and evaluated. The results are summarized in Table 9. According to this analysis, mean grades assigned are not very different in honors sections from those in other sections, except in the University (Basic) College. In the University College, however, grading practices are significantly higher and there are more honors sections in this college than in any other. This state of affairs in University College influences the total interpretation of grading practices in honors sections. The mean grades assigned in honors sections of the Business College are actually lower than those assigned in other courses, although not significantly so.

TABLE 8. Matched-Pairs Signed-Ranks Test of Average Grades Assigned in Courses in Which Common Final Examinations Were and Were Not Administered

Discipline	Level	Common Finals Average Diff. in Mean Grades	Other Than Common Finals Average Diff.	Difference	Absolute Rank of Difference	R(+)	R(-)
Language	L.D.	-.60	-.48	.12	3.5	3.5	
Language	L.D.	-.05	.22	.27	9	9	
Prof.	L.D.	-.02	.26	.28	10.5	10.5	
Prof.	U.D.	-.34	.27	.61	15	15	
Science	L.D.	.15	-.41	-.56	14		14
Science	L.D.	.12	-.52	-.64	16		16
Science	L.D.	-.50	-.28	.22	5	5	
Science	U.D.	-.69	-.43	.26	7.5	7.5	
Science	U.D.	-.36	-.37	-.01	1		1
Science	L.D.	-.32	-.20	.12	3.5	3.5	
Science	L.D.	-.15	-.40	-.25	6		6
Science	L.D.	-.32	-.06	.26	7.5	7.5	
Science	L.D.	-.45	-.10	.35	12	12	
Science	L.D.	-.16	.12	.28	10.5	10.5	
Science	L.D.	-.24	.26	.50	13	13	
Soc. Sci.	L.D.	-.18	-.26	-.08	2		2
Median		-.28	-.23	.24		97	T = 39

A table value of 30 or less required for a two-tailed test at the .01 level (Downie & Heath, 1965)

TABLE 9. Differences in Mean Grades Assigned in Honors Sections and All Other Courses Enrolled by the Same Students at Michigan State University, Fall, 1965

College	Number of Honors Sections	Differences In Mean Grades	t-ratio	Table Value*
Arts & Letters	20	.20	1.03	2.86
Business	5	-.20	-1.65	4.60
Education	2	.42	7.69	63.66
Home Economics	2	.02	.38	63.66
Natural Science	16	.06	.39	2.95
Social Science	10	.18	1.58	3.25
Sub-Total	55	.12	1.33	2.68
University (Basic)	30	.28	5.19*	2.75
Total University	85	.18	2.82*	2.64

*Significant at the .01 level

To investigate the relationship between section size and average grades, Pearson Product-Moment Coefficients of Correlation between the two variables were calculated for all courses with 20 or more sections. Honors sections in the University College were excluded from this analysis because it has been determined that grading practices are quite different in those sections than they are in regular sections. The correlations are presented in Table 10. There seems to be little relationship between section size and average grades assigned to a section. Ten of the courses had small, insignificant negative correlations. Of the 14 positive correlations, 3 were significant.

TABLE 10. Correlation of Section Size and Average Grades in Courses with 20 or More Sections at Michigan State University, Fall, 1965

Course	Number of Sections	r	Average Section Size	Average Grade
1	133	-.004	49	2.26
2	91	.006	49	2.36
3	146	.271*	40	2.27
4	20	.006	37	2.33
5	68	-.019	52	2.37
6	33	.006	15	2.09
7	30	.042	27	2.62
8	84	-.082	6	2.79
9	23	.005	19	2.24
10	21	-.336	18	2.13
11	47	.359*	17	2.03
12	23	.005	20	2.74
13	35	-.136	22	2.29
14	26	-.222	33	2.52
15	23	-.268	26	2.35
16	157	.165*	39	2.54
17	22	.039	37	2.83
18	44	-.055	23	2.33
19	72	.000	22	1.91
20	31	.276	32	2.02
21	44	.000	32	2.21
22	28	-.170	26	2.22
23	23	.240	28	2.23
24	32	-.280	21	2.17

* Significant at the .05 level

To gain information about grading practices in remedial courses, the data in Table 11 were taken directly from the "Grading Practices Analysis" report for Fall, 1965. The data presented in Table 11 indicate quite conclusively that the mean grades assigned in these remedial sections are different from grades assigned to the same students in all of the other courses in which they were enrolled at the same time. In two of the courses the average grades assigned were considerably higher; in the third they were extremely low.

Discussion

Differences in mean grades assigned were found among, as well as within, colleges, levels, and departments. Variability of mean grades assigned, indicated by analysis of variance of difference in mean grades, was found among colleges but not among levels. Hence, average grades increase with level; but variability within colleges remains fairly constant through all levels. Interaction between colleges and levels was insignificant for sample mean grades, but when differences in mean grades were considered there was significant interaction. Thus, ability levels may be different from college to college, as well as the recognized difference in achievement from level to level. The increased sensitivity of the difference in mean grades index over the mean grade in section may also be reflected.

TABLE 11. Difference in Grading Practices in Remedial Courses at Michigan State University, Fall, 1965

Course	Number of Grades in Section	Mean Grade in Section	Number of Grades in Other Courses	Mean Grade in Other Courses	Difference in Mean Grades	T-Test of Difference*
Arith. Impr. Ser.	33	3.09	121	1.91	1.18	9.52*
	32	2.81	113	1.37	1.44	9.45*
	33	2.88	118	1.91	0.97	7.83*
	32	3.31	121	1.99	1.32	9.81*
Total	130	3.02	473	1.80	1.22	17.79*
Algebra	231	0.67	797	2.00	-1.33	-20.85*
	179	0.51	599	2.01	-1.50	-18.52*
Total	410	0.60	1396	2.00	-1.40	-27.56*
Methods of Study	25	2.80	83	1.78	1.02	4.29*
	23	2.52	65	1.96	0.56	2.66
	22	3.00	66	2.16	0.84	3.89*
	23	2.91	75	1.88	1.03	4.69*
	19	2.79	69	2.06	0.73	3.10*
	22	3.05	76	2.19	0.86	4.81*
	21	2.62	71	1.93	0.69	2.97*
	21	3.24	74	1.86	1.38	5.81*
	23	3.00	73	1.87	1.13	4.35*
	20	2.70	68	1.94	0.76	3.40*
	23	2.96	84	2.10	0.86	4.66*
	22	3.05	79	1.84	1.21	4.30*
Total	264	2.89	883	1.96	0.93	14.00*

*Significant at the .01 level

Common final examinations lessen the variability of grading practices while not affecting average grading. This practice appears to make a positive contribution toward consistency as compared with inconsistencies of grading practices observed elsewhere in the survey. Of course the use of common final examinations is limited to large-enrollment courses with several sections of the same course offered during a given term.

Grading practices in honors sections were not very different from those in regular courses except in the University (Basic) College. To propose the proper level of average grades is not our purpose, but certainly it should not be expected that students in an honors section would be graded lower or higher on the average than in all other courses taken at the same time.

Observations regarding relationships between section size and average grades must be tempered in that the correlations were low and balanced as far as negative and positive directions are concerned. The slight tendency toward positive relationships might be interpreted to mean that as class size increases it is impossible to accurately judge the degree of attainment of objectives of a course, and the students are given the benefit of any doubt.

The disparity of average grades assigned in remedial courses might be attributed to the fact that a student enrolls in a course such as this because of a particular weakness. The relatively low mean grades in other courses indicates that students enrolled in remedial courses persist with a low achievement level in other courses, which results in large and significant differences in mean grades.

Conclusions

One of the objectives of institutional research is to offer, after careful study of a problem, some alternative courses of action. One alternative to the problems of grading might be to expand the use of a pass-fail system of marking. This was done at the University of Chicago in 1931 (Reeves, et al , 1933) and is tried at many other institutions of higher education. Usually a pass-fail course must be taken outside a student's major field, and at some colleges they do not count toward graduation requirements (Aiken,1967). The possible outcomes include reduced pressure on the student and teacher, but it is not clear whether this would result in an increased or decreased level of achievement. For various reasons registrars, scholarship directors, graduate schools, athletic directors, parents, and even students have called for a return to the familiar A through F grading system. The University of Chicago (1965a, 1965b) has returned to the traditional A through F system in conjunction with a High Pass = 4 and Pass = 2. As Dressel (1954) described the situation: "Students would like all A's, administrators would like few F's, and the teachers would like to be left alone."

Another alternative that would leave the teachers alone, at least for a while, is likely to consternate the registrars. This system would provide for the correction of grading inequities through the application of an "index of grading." This is merely the ration of the mean grade in section to the mean grade in other courses. It was first suggested by Reeves, et al, (1933) before the advent of the computer. The data in the "Grading Practices Analysis" report provide all that is necessary to program for what could be called an "All University Grade Point Average." (AUGPA) A section which is assigned higher grades than its students earn in other courses would have an index of grading greater than 1.00; one which is assigned lower grades than the same students earn in other courses would have an index of grading of less than 1.00. The AUGPA would be weighted accordingly with the index of grading.

The benefits of this alternative treatment of grades are many. Students would not be evaluated by a few professors, some of whom may be very subjective in their judgments and prone to be extremely harsh or overly generous in their grading. Each student's AUGPA would be determined by the grading practices of the professors of the entire university. It would not be necessary for a student to switch majors to graduate or to maintain a respectable grade point average. On the other hand, there would be no havens for the less able or unmotivated students to while away their time. This procedure of correcting grades would provide for an evaluation of the theory of correlated reinforcement. Correlated reinforcement can be defined as a condition when the quanity of satisfaction is related in some way to performance. Travers (1967) suggests: "An example of correlated reinforcement is the assignment of a grade so that the grade reflects the quality of performance." Many of the possible benefits of correlated reinforcement are currently lost through inconsistent grading practices.

Limitations of this alternative procedure also exist. Only grade point averages have been considered, and the critical points on a grading scale exist at the upper and lower end of the continuum. This area is where a student is considered for either academic honors or dismissal, and this limitation was also recognized by Reeves, et al, (1933). It might also be argued that if an able student selects as his major a department where high grading practices are prevalent, an earned A may ultimately be evaluated as something less than a 4.0. A less able student may suffer by choosing courses or being assigned to a section where an earned C is similarly affected. The most damaging limitation of all may be that special abilities required for success in one course are very different from those required in another. If this is true, then it is unrealistic to expect grading indices to approach 1.0. In response to this, Reeves, et al,(1933) stated: "The fact that academic groups are willing to grant a common degree, the bachelor's for attainment of very different types possibly amounts to a recognition that there is a central core of common elements that is more important than the elements peculiar to specialization in certain fields." If we are willing to grant a common degree, then perhaps we should at least consider the alternative of producing common grading practices.

A less drastic proposal for attempts to reduce the variability in grading standards observed in this study is to encourage instructors to share information and rationale on the grades they assign their students. In colleges and universities in which grades are the private property of the instructor and the student, variation in grading practices and standards should be expected. The availability of information such as that contained in the "Grading Practices Analysis" report on which this study was based provides feedback to the instructor that can be very useful to him. Information on grade distributions for sections, courses, and departments, along with encouragement to faculty members to review and discuss it, can be expected to serve as a self-correcting technique for reducing the variability in grading practices frequently observed.

References

Aiken, D.L., "'Pass-fail' Grading Option Adopted on Many Campuses," *The Chronicle of Higher Education*, I (1967), 1.

American Psychological Association, Council of Editors, *Publication Manual of the American Psychological Association*. Washington, 1957.

Bloom, B.S., (ed.) *Taxonomy of Education Objectives: The Classification of Educational Goals*. New York, 1956.

Downie, N.M. and Heath, R.W., *Basic Statistical Methods*. New York, 1965.

Dressel, P.L., (ed.) *Evaluation in General Education*. Dubuque, Iowa, 1954.

Juola, A.E., *Illustrative Problems in College-Level Grading*. Paper at a joint session of the American Educational Research Association and National Council on Measurement in Education, New York, February, 1967.

Petersen, Dwain F., "The University of Chicago Graduate Programs in the Divisions," LXVI, *Announcements 1966-67*. Chicago, 1965.

Peterson, Dwain F., *The University of Chicago Student Handbook 1965-67*. Chicago, 1965.

Reeves, F.W., Peik, W.E., and Russell, J.D., *The University of Chicago Survey*. Vol. IV. *Instructional Problems in the University*. Chicago, 1933.

Travers, R.M.W., *Essentials of Learning*. New York, 1967.